UNCLASSIFIED

AFOSR 2036

SYNTHESIS OF SHAPING FILTERS FOR NONSTATIONARY STOCHASTIC PROCESSES AND THEIR USES

Edwin Stear

(Report No. 61-50)
University of California
Department of Engineering
Los Angeles, California

August 1961

CONTRACT NO. AF 49(638)438

MECHANICS DIVISION
AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
AIR RESEARCH AND DEVELOPMENT COMMAND
WASHINGTON 25, D.C.

AFOSR 2036

SYNTHESIS OF SHAPING FILTERS FOR NONSTATIONARY STOCHASTIC PROCESSES AND THEIR USES

Edwin Stear

(Report No. 61-50)
University of California
Department of Engineering
Los Angeles, California

August 1961

CONTRACT NO. AF 49(638)438

MECHANICS DIVISION
AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
AIR RESEARCH AND DEVELOPMENT COMMAND
WASHINGTON 25, D.C.

FOREWORD

The research described in this report, *Synthesis of Shaping Filters for Nonstationary Stochastic Processes and Their Uses,* by Edwin Stear was carried out under the technical direction of C. T. Leondes and Gerald Estrin and is part of the continuing program in Adaptive Control Systems.

This project is conducted under the sponsorship of the Air Force Office of Scientific Research of the Air Research and Development Command.

Submitted in partial fulfillment of
Contract Number Af 49(638)438
Task Number 9783

J. M. English
Director - Engineering Research

UNIVERSITY OF CALIFORNIA
DEPARTMENT OF ENGINEERING
LOS ANGELES, CALIFORNIA

ACKNOWLEDGMENTS

The author gratefully acknowledges the constant assistance and encouragement so freely given him by his Doctoral Committee. He shall be forever indebted to Professor C. T. Leondes for the opportunity of studying under him and working for him as a student and a member of the Adaptive Control Systems Research Group at UCLA. The author also acknowledges his great indebtedness to Professor H. Davis whose wise counsel was responsible for the decision by the author to study for a Ph.D. in Engineering and whose stimulating informal discussions played so large a part in making its completion possible. The author thanks Professor T. Ferguson for his patience with and assistance to an engineer struggling with some of the more intricate aspects of mathematical statistics and stochastic processes and Professor M. Aoki for the many discussions which contributed much to the author's knowledge of control systems theory. He would especially like to thank Professor E. Beckenbach for agreeing to serve on his Doctoral Committee on very short notice and for so willingly bearing the inconvenience it entailed.

The author thanks Messrs. Stubberud, Hsieh, Nesbit, Schwartz, and Bekey for the privilege of working with them and learning from them as a fellow member of the Adaptive Control Systems Research Group. He also gratefully acknowledges the contribution of Mr. Al Stubberud to his knowledge of linear systems theory and for the privilege of collaborating with him on the work in Chapters 2 and 3 of this dissertation and that published elsewhere.

ABSTRACT

This dissertation considers the problem (usually called the shaping filter problem) of synthesizing linear systems whose responses to *white noise* input processes and appropriate sets of random initial conditions will be stochastic processes whose covariance functions are prescribed functions of two variables.

A detailed review of previous work on this problem is presented and the limitations of and the errors in this previous work are carefully pointed out. The properties of weighting functions of systems characterizable by finite-order linear differential equations are developed in detail and these results are used to develop the properties of the covariance functions of the responses of such systems to *white noise* inputs and random initial conditions.

Based on this work, exact solutions to the shaping filter problem are presented for certain special cases and some discussion of predictability of the processes is presented. Attention is then focused on the general problem and by means of the Schauder Fixed Point Theorem and by Picard's method of successive approximations the existence of physically realizable shaping filters is established for a large class of separable covariance functions. The question of the uniqueness of the shaping filter and its relationship to the covariance matrix of the set of random initial conditions is investigated. It is shown that if an appropriate set of random initial conditions is specified, then the weighting function of the shaping filter is unique up to a multiplicative factor of ± 1. The further requirements on the covariance function in order to guarantee that the shaping filter can be characterized by a finite-order, linear differential equation with continuous coefficients are given as well as certain lesser requirements which permit easy analog simulation.

Some brief comments are made relative to computational requirements and how they can best be carried out and the pertinent references are cited. Finally the two main areas of applications of shaping filters are briefly outlined.

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

1.1 A GENERAL DESCRIPTION OF THE PROBLEM

A problem of some current interest in the field of stochastic processes can be described briefly as follows, (see Figure 1).

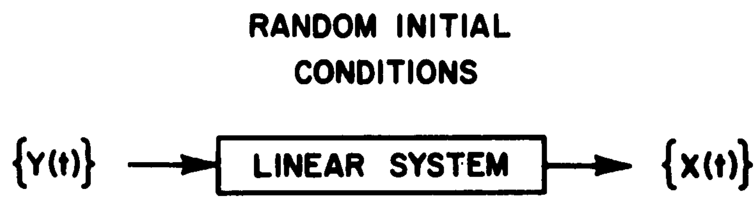

FIGURE 1

Given the covariance function $\Gamma(t_1,t_2)$ of some continuous in the mean stochastic process, determine the weighting function and/or differential equation, if appropriate, of a continuous, physically realizable, linear system and the covariances of a set of random initial conditions, if required, such that the response $\{X(t)\}$ of the system to a *white noise* input process $\{Y(t)\}$ and the set of random initial conditions will be a stochastic process whose covariance function $\Gamma_{XX}(t_1,t_2)$ satisfies the equality $\Gamma_{XX}(t_1,t_2) = \Gamma(t_1,t_2)$. This problem which will be referred to from hereon as the shaping filter problem,* has not yet been solved for the general case where $\Gamma(t_1,t_2)$ is an arbitrary continuous covariance function; i.e., an arbitrary, bounded, nonnegative-definite, continuous function defined on a region $T \times T$ of the real plane where T is an interval of the real line. However, by suitably restricting the class of admissible covariance functions and the interval T, certain fairly definitive results have been obtained. These results are summarized in the following section.

This dissertation is devoted to a study of the shaping filter problem for the class of separable covariance functions, where T is a finite or semi-infinite (to the right) interval, and to the relationship between the properties of the covariance function and the possibility of characterizing the shaping filter by a finite-order, linear differential equation. Particular emphasis is placed on the goal of obtaining the solution to the problem in a form

*There is, of course, a related problem for discrete parameter stochastic processes; but, since these two problems are quite closely related, attention herein is directed primarily to the continuous parameter case.

directly applicable in engineering practice, as is proper for a dissertation in engineering. Finally, the applications of the solution to the shaping filter problem to the problems of analog simulation and linear, least-squares filtering is discussed.

1.2 HISTORICAL REVIEW OF THE SHAPING FILTER PROBLEM

In order to place the results obtained in this dissertation in proper perspective, it is necessary to have a precise summary of all significant previously published works on the shaping filter problem. However, before such a summary is given, a brief chronological sketch of this work is in order.

The first significant result relative to the shaping filter problem was apparently obtained by Wold [1,1938] for discrete parameter stationary* processes and is part of his fundamental decomposition theorem [1,p.89]. Kolmogorov [2,1939; 3,1941; 4,1941] then put Wold's decomposition theorem in an analytic setting and obtained some new theorems for discrete parameter stationary processes, parts of which again pertain to the shaping filter problem. Independently, Wiener [5,1942] obtained Kolmogorov's results for discrete parameter stationary processes with absolutely continuous spectral distribution functions and generalized the results to include continuous parameter stationary processes with absolutely continuous spectral distribution functions. Wiener obtained thereby, as part of his results, the first solution of the shaping filter problem for continuous parameter stationary processes. Then Hanner [6,1949] and Karhumen [7,1950] obtained, by different methods, the continuous parameter analog of the Wold decomposition theorem, parts of which again pertain to the shaping filter problem. Bode and Shannon [8,1950], in their simplified heuristic derivation of Wiener's results on linear, least-squares, prediction and filtering theory stressed the solution of the shaping filter problem as an important step in their method. The next significant result relative to the shaping filter problem was obtained by Darlington [9,1959] in his generalization of Bode and Shannon's work so as to include nonstationary processes. While Darlington did make a good start on the problem for continuous parameter nonstationary processes, he did not obtain very useful results as will be clear from later discussion. Two months later, Batkov [10,1959] published a paper which presents three methods for solving the shaping filter problem, including an *algebraic* method using various partial derivatives of the covariance function, for a certain class of continuous parameter nonstationary processes. As will be shown later, Batkov's algebraic procedure only works for a rather

---

**Stationary* should always be interpreted as "wide-sense stationary".

specialized subclass of the class claimed. Later in the same year, Sondhi and Higgins [11,1959] presented a solution to a modified form of the shaping filter problem requiring the use of several *white noise* sources. Nothing more will be said of this work because interest in this dissertation is restricted to the use of one *white noise* input process. Finally, Leonov [12,1960] presented a rather nice mathematical solution to the shaping filter problem for continuous parameter processes (both stationary and nonstationary) in terms of expansions in orthogonal functions.

In the following paragraphs a precise critical summary of the above mentioned work will be given.

1.2.1 Summary of Previous Work for Stationary Processes

Since no useful purpose would be served by it insofar as this dissertation is concerned, no attempt will be made to summarize the previous work for stationary processes individually as cited in the chronological sketch. Rather, an over-all summary will be given, the details of which can be found in the books by Doob [13,pp.527-559, 569-590] and Grenander and Rosenblatt [14,pp.65-82]. Only the continuous case is considered and, naturally, $T = (-\infty, \infty)$.

If the stochastic process $\{Y(t)\}$* is stationary and continuous in the mean, then it has the spectral representation**

$$Y(t) = \int_{-\infty}^{\infty} e^{i2\pi\lambda t} dZ(\lambda) \tag{1-1}$$

where the process $\{Z(\lambda)\}$ has orthogonal increments and $E|dZ(\lambda)|^2 = dF(\lambda)$. $F(\lambda)$ is called the spectral distribution function of $\{Y(t)\}$ and

$$\Gamma_{YY}(\tau) = E\, Y(t)\, Y(t+\tau) = \int_{-\infty}^{\infty} e^{i2\pi\lambda\tau}\, dF(\lambda) \tag{1-2}$$

Furthermore, $F(\lambda)$ is nondecreasing and since

$$\int_{-\infty}^{\infty} dF(\lambda) = \Gamma_{YY}(0) < \infty \tag{1-3}$$

*$\{Y(t)\}$ is assumed to be real valued.

**For a definition of all stochastic integrals herein see Reference 13, p. 426.

$F(\lambda)$ is also of bounded variation. Hence, $F(\lambda)$ can be decomposed into the sum of three nondecreasing functions

$$F(\lambda) = F_1(\lambda) + F_2(\lambda) + F_3(\lambda) \tag{1-4}$$

where $F_1(\lambda)$ is the jump function part of $F(\lambda)$, $F_2(\lambda)$ is the absolutely continuous part of $F(\lambda)$, and $F_3(\lambda)$ is the continuous singular part of $F(\lambda)$. This decomposition of $F(\lambda)$ corresponds to a decomposition of $\{Y(t)\}$ into three mutually orthogonal processes $\{Y_1(t)\}$, $\{Y_2(t)\}$, and $\{Y_3(t)\}$ with spectral distribution functions $F_1(\lambda)$, $F_2(\lambda)$ and $F_3(\lambda)$ respectively.

If $\{Y(t)\}$ is applied to the input of a linear system whose frequency response function,* $G(\lambda)$, satisfies the condition

$$\int_{-\infty}^{\infty} |G(\lambda)|^2 \, dF(\lambda) < \infty \tag{1-5}$$

then the system output, $\{X(t)\}$, will be a continuous in the mean, stationary process whose spectral distribution function, $F_X(\lambda)$, is given by

$$F_X(\lambda) = \int_{-\infty}^{\lambda} |G(\lambda)|^2 \, dF(\lambda) \tag{1-6}$$

It should be noted that $G(\lambda)$ is not, in general, required to be in L_2 (i.e., it is not required that $\int_{-\infty}^{\lambda} |G(\lambda)|^2 \, d\lambda < \infty$). From (1-6) it follows that $F_X(\lambda)$ is absolutely continuous if $F(\lambda)$ is. If $F(\lambda)$ is absolutely continuous and if $|f(\lambda)|^2 = F'(\lambda)$, then (1-1) can be replaced by

$$Y(t) = \int_{-\infty}^{\infty} e^{i2\pi\lambda t} f(\lambda) \, d\tilde{Z}(\lambda) \tag{1-7}$$

where $\{\tilde{Z}(\lambda)\}$ has orthogonal increments and $E|d\tilde{Z}(\lambda)|^2 = d\lambda$. On the other hand suppose $\{Y(t)\}$ is generated from a process $\{V(t)\}$ according to the equation

$$Y(t) = \int_{-\infty}^{\infty} W(\tau) \, dV(t-\tau) \tag{1-8}$$

where $\{V(t)\}$ has orthogonal increments with $E|dV(t)|^2 = dt$ and $\int_{-\infty}^{\infty} |W(\tau)|^2 d\tau < \infty$. Then

$$\Gamma_{YY}(t_1, t_2) = \int_{-\infty}^{\infty} W(t_1-\theta) \, W(t_2-\theta) \, d\theta \tag{1-9}$$

---

* $G(\lambda)$ is Doob's gain function $C(\lambda)$.

From (1-9) it follows that $\{Y(t)\}$ is stationary and continuous in the mean, Furthermore

$$F(\lambda) = \int_{-\infty}^{\lambda} |G(\lambda)|^2 \, d\lambda \leq \int_{-\infty}^{\infty} |G(\lambda)|^2 \, d\lambda = \int_{-\infty}^{\infty} |W(\tau)|^2 \, d\tau < \infty \qquad (1\text{-}10)$$

where $G(\lambda)$ is the Fourier Transform of $W(\tau)$ and, hence, $F(\lambda)$ is absolutely continuous and $F'(\lambda) = |G(\lambda)|^2$. Formally considering the increments of $\{V(t)\}$ to be given by

$$V(t_2) - V(t_1) = \int_{t_1}^{t_2} U(t) \, dt \qquad (1\text{-}11)$$

where $\{U(t)\}$ is a *white noise* process, (1-8) represents the response of a linear system with weighting function, $W(\tau)$, to a *white noise* input process and (1-9) and (1-10) represent well known results which are usually obtained in a less rigorous way by engineers.*

As a consequence of the above results, a simple necessary and sufficient condition for the existence of a solution to the shaping filter problem for continuous in the mean, stationary processes can be stated, providing the requirement of physical realizability is waived. They are: If $F(\lambda)$ is the spectral distribution function corresponding to the given covariance function, $\Gamma(\tau)$, then a process, $\{Y(t)\}$, whose covariance function $\Gamma_Y(\tau)$ satisfies the equality $\Gamma_Y(\tau) = \Gamma(\tau)$ can be generated from a *white noise* process by means of a linear system if and only if $F(\lambda)$ is absolutely continuous. Moreover, any linear system whose frequency response function, $G(\lambda)$, satisfies the equality $|G(\lambda)|^2 = F'(\lambda)$ almost everywhere can be used to generate such a process, $\{Y(t)\}$. More generally, even if $F(\lambda)$ is not absolutely continuous, the above still applies to the absolutely continuous part of $F(\lambda)$; i.e., to $F_2(\lambda)$ in the decomposition given above. If the requirement of physical realizability is not waived, then the above condition must be strengthened somewhat. Many years ago, Paley and Wiener [15, p.16, Theorem XII] showed that if $\int_{-\infty}^{\infty} |G(\lambda)|^2 \, d\lambda < \infty$, where $G(\lambda)$ is the frequency response function of a linear system, then the system is physically realizable if and only if**

$$\int_{-\infty}^{\infty} \frac{|\log G(\lambda)^2|}{1+\lambda^2} \, d\lambda < \infty \qquad (1\text{-}12)$$

In view of the required equality $|G(\lambda)|^2 = F'(\lambda)$ i.e., physical realizability of the linear system (the shaping filter) requires the additional condition

*The usual engineering procedure could perhaps be rigorized at the expense of introducing generalized linear functionals.

**The assumed continuity of $\Gamma(\tau)$ guarantees that $\int_{-\infty}^{\infty} |G(\lambda)|^2 \, d\lambda < \infty$.

$$\int_{-\infty}^{\infty} \frac{|\log F'(\lambda)|}{1+\lambda^2}\, d\lambda < \infty \tag{1-13}$$

i.e., for physical realizability of the shaping filter, $F(\lambda)$ must be absolutely continuous and satisfy (1-13). Since the above conditions depended only on the magnitude of $G(\lambda)$, it is clear that these conditions do not uniquely determine the shaping filter. A desirable* way of rendering the shaping filter essentially unique (to determine the weighting function uniquely except on a set of Lebesque measure zero) is to require it to be a minimum phase filter; i.e., to require that $G(\lambda) \neq 0$ for $Im\ \lambda < 0$. Such a $G(\lambda)$ is given by the (loss-phase) integral

$$G(\lambda) = \exp\left[-\frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{(1+\lambda\omega)\log F'(\omega)}{(\lambda-\omega)(1+\omega^2)} \right] \tag{1-14}$$

Except for some brief comments relative to some special cases of stationary processes which appear at appropriate places throughout the remainder of this dissertation, this concludes the summary of previous work for stationary processes. Clearly, for continuous in the mean, stationary processes and $T = (-\infty, \infty)$ the shaping filter problem had been resolved in rather definitive terms prior to 1950.**

1.2.2 Summary of the Work of Darlington

In his paper [9], Darlington presents a generalization of Bode and Shannon's results [8] so as to include nonstationary processes. In his generalization, as in Bode and Shannon's original procedure, the central problem, of course, is that of resolving the shaping filter problem. As will be clear from the summary to follow, the shaping filter problem considered by Darlington is somewhat different from (but related to) that described in Section 1.1.

After some preliminary remarks on the Bode-Shannon model and its use of shaping filters, Darlington turns his attention to the shaping filter problem, proceeding as follows. If $W(t,\tau)$ denotes the weighting function of a linear system then the covariance function $\Gamma(t_1, t_2)$ of the output of the system, when the input is a *white noise* process, is given, when it exists, by the expression

$$\Gamma(t_1, t_2) = \int_{-\infty}^{\infty} W(t_1, \tau)\, W(t_2, \tau)\, d\tau \tag{1-15}$$

Note that, since the lower limit of the integral is $-\infty$, it has tacitly been assumed that the *white noise* input has been applied to the system continuously

*The inverse filter corresponding to a physically realizable, minimum phase filter is physically realizable and stable. This is important for applications to linear, least-squares filtering and prediction theory.

**The shaping filter problem is apparently still unresolved for stationary processes which are not continuous in the mean.

throughout the infinite past. If the system is physically realizable, then $W(t, \tau) = 0$ for $\tau > t$ and the upper limit of the integral in (1-15) can be replaced by $\min[t_1, t_2]$. Letting $W^a(t, \tau)$ denote the weighting function of the adjoint system, $W^a(t,\tau) = W(\tau, t)$, and $\Gamma(t_1, t_2)$ can be expressed in the equivalent form

$$\Gamma(t_1, t_2) = \int_{-\infty}^{\infty} W(t_1,\tau)W^a(\tau, t_2)d\tau \tag{1-16}$$

Since (1-16) expresses $\Gamma(t_1, t_2)$ as the convolution of two weighting functions, $\Gamma(t_1, t_2)$ can be interpreted as the weighting function of the nonphysically realizable (self-adjoint) system composed of the original system in cascade with its corresponding adjoint system.

Because Darlington was unable to find a suitable nonstationary analog of the loss-phase integral (i.e., Equation 1-14) for solving the nonstationary shaping filter problem, he restricts his attention to systems which are completely characterized by finite-order, linear differential equations and seeks an analog, in terms of operations with differential equations, of the usual procedure of factorization of the rational spectral density function in the corresponding stationary case.

When the system is completely characterized by a finite-order linear differential equation,* then its response V is related to its excitation E by an expression of the form

$$B(p, t)\, V(t) = H(t)\, A(p, t)\, E(t) \tag{1-17}$$

where $B(p, t)$ and $A(p, t)$ are polynomials in p with time-varying coefficients where $p = \frac{d}{dt}$; i.e.,

$$\begin{aligned} B(p, t) &= p^n + b_{n-1}(t)p^{n-1} + \ldots + b_0(t) \\ A(p, t) &= p^m + a_{m-1}(t)p^{m-1} + \ldots + a_0(t) \end{aligned} \tag{1-18}$$

and $H(t)$ is a time-varying scale factor. Any set of n linearly independent solutions, say $U_i(t)$, $i = 1, \ldots, n$, of

$$B(p, t)\, V(t) = 0 \tag{1-19}$$

*See Chapters 2 and 3 of this dissertation for details of some of the following discussion.

form a set of basis functions (*bf's*) for $B(p, t)$ and for the system. Similarly, any set of m linearly independent solutions of

$$A(p, t)\, E(t) = 0 \tag{1-20}$$

form a set of basis functions for $A(p, t)$ and are called the zero response functions (*zrf's*) of the system. If the system is also time-invariant (stationary), then the *bf's* and *zrf's* are exponentials,* $e^{S_\sigma t}$ where the S_σ's are the familiar poles and zero of the system transfer function. The *bf's* and *zrf's* of non-stationary systems play analogous, equally important roles even when they cannot be represented by simple coefficients like S_σ.

When two systems are cascaded where both are completely characterized by a finite-order differential equation, then the over-all system is completely characterized by a finite-order differential equation corresponding to the *product* of the differential equations of the two given systems. In terms of operators, the *product* may be represented by **

$$B_1 V_1 = H_1 A_1 E, \; B_2 V = H_2 A_2 V_1 \;, \; BV = HAE \tag{1-21}$$

where $BV = HAE$ is the differential equation of the over-all system. As Darlington indicates, the operators B, A, and H can be determined from B_1, B_2, A_1, A_2, H_1, and H_2 by means of derivative and algebraic operations. This corresponds formally to the convolution of the weighting functions of the two given systems. Similarly, corresponding formally to the sum of weighting functions is a suitably defined *sum* of their corresponding differential equations represented by

$$\begin{gathered} B_1 V_1 = H_1 A_1 E \;, \quad B_2 V_2 = H_2 A_2 E \\ V = V_1 + V_2 \;, \qquad BV = HAE \end{gathered} \tag{1-22}$$

The operators B and A and the scale factor H can also be determined from B_1, B_2, A_1, A_2, H_1, and H_2 by means of derivative and algebraic operations. Further, the *bf's* of B are those of B_1 plus those of B_2, but the *bf's* of A (the *zrf's* of the *sum* system) are not related to those of A_1 and A_2 in any simple way.

Corresponding to the system characterized by (1-17) is its related adjoint system which is completely characterized by the adjoint differential equation

$$B^a(p, t)\, V(t) = \pm\, H(t)\, A^a(p, t)\, E(t) \tag{1-23}$$

---

*Or linear combinations of them.

**Suppressing the arguments for convenience of notation.

corresponding to (1-17), the operators $B^a(p, t)$ and $A^a(p, t)$ being easily determined from $B(p, t)$ and $A(p, t)$. When the system is physically realizable, the weighting function corresponding to (1-17) can be expressed in the form*

$$W(t, \tau) = \begin{cases} \sum_{i=1}^{n} \frac{U_i(t)}{U_i(\tau)} J_i(\tau) & , t > \tau \\ 0 & , t < \tau \end{cases} \tag{1-24}$$

and that corresponding to (1-23); i.e., that of the nonphysically realizable adjoint system; in the form

$$W^a(t,\tau) = \begin{cases} \sum_{i=1}^{n} \frac{U_i(\tau)}{U_i(t)} J_i(t) & , t < \tau \\ 0 & , t > \tau \end{cases} \tag{1-25}$$

The *product* of (1-17) and (1-23) corresponds to the convolution of $W(t, \tau)$ and $W^a(t, \tau)$ as in (1-16) and is written as

$$B^{\Gamma}(p, t)\, V(t) = \pm H^2(t)\, A^{\Gamma}(p, t)\, E(t) \tag{1-26}$$

From the discussion following (1-16), it is clear that the weighting function of the system characterized by (1-26) is $\Gamma(t_1, t_2)$, which, from (1-16), (1-24), and (1-25), can be written in the form

$$\Gamma(t_1, t_2) = \begin{cases} \sum_{i=1}^{n} \frac{U_i(t_1)}{U_i(t_2)} Q_i(t_2) & , t_1 > t_2 \\ \sum_{i=1}^{n} \frac{U_i(t_2)}{U_i(t_1)} Q_i(t_1) & , t_1 < t_2 \end{cases} \tag{1-27}$$

The symmetry of $\Gamma(t_1, t_2)$ expresses the fact that (1-26) is a self-adjoint equation.

With this background, Darlington takes up the shaping filter problem as encountered in the Bode-Shannon model, assuming that the *signal, S(t),* and *noise, N(t),* are generated from uncorrelated *white noise* sources by means of physically realizable systems characterized by finite-order, linear differential equations. If the *bf's* and *zrf's* of the systems are known, then the weighting functions of the systems, $W_S(t,\tau)$ and $W_N(t,\tau)$, are easily determined and

*Assuming order n of B > order m of A.

$\Gamma_S(t_1, t_2)$ and $\Gamma_N(t_1, t_2)$ found from (1-16). Also, differential equations of the form in (1-26) can be found for both $\Gamma_S(t_1, t_2)$ and $\Gamma_N(t_1, t_2)$ by forming the *product* of the corresponding differential equations and their adjoints as described above (even if the *bf's* and *zrf's* of the generating systems are unknown). If $F = S + N$, then $\Gamma_F(t_1, t_2) = \Gamma_S(t_1, t_2) + \Gamma_N(t_1, t_2)$ and a differential equation of the form in (1-26) whose corresponding weighting function is $\Gamma_F(t_1, t_2)$ can be found from $\Gamma_F(t_1, t_2)$ itself or by *summing* the differential equations corresponding to $\Gamma_S(t_1, t_2)$ and $\Gamma_N(t_1, t_2)$ in case $\Gamma_F(t_1, t_2)$ is unknown. In this way, there is determined

$$
\begin{aligned}
B_S^{\Gamma}(p,t)\ V(t) &= \pm\ H_S^2(t)\ A_S^{\Gamma}(p,t)\ E(t) \\
B_N^{\Gamma}(p,t)\ V(t) &= \pm\ H_N^2(t)\ A_N^{\Gamma}(p,t)\ E(t) \qquad (1\text{-}28) \\
B_F^{\Gamma}(p,t)\ V(t) &= \pm\ H_F^2(t)\ A_F^{\Gamma}(p,t)\ E(t)
\end{aligned}
$$

The *bf's* of $B_F^{\Gamma}(p,t)$ are those of $B_S^{\Gamma}(p,t)$ and $B_N^{\Gamma}(p,t)$ and are even in number, one-half of them being the *bf's* of the systems used to generate $S(t)$ and $N(t)$ and the other half being the *bf's* of the corresponding nonphysically realizable adjoint systems. On the other hand, the *bf's* of $A_F^{\Gamma}(p,t)$, again even in number, are not simply related to the *bf's* of $A_S^{\Gamma}(p,t)$ and $A_N^{\Gamma}(p,t)$ and must be found as the solutions as

$$A_F^{\Gamma}(p,t)\ E(t) = 0 \qquad (1\text{-}29)$$

This corresponds to the calculation of the zeros of the rational signal-plus-noise spectral density function in the stationary case in which the spectral densities of S and N are added (corresponding to forming the sum of the differential equations for $\Gamma_S(t_1, t_2)$ and $\Gamma_N(t_1, t_2)$) to get the spectral density of F. The addition retains the poles but the zeros must be calculated as the zeros of a polynomial (corresponding to finding the solutions of Equation 1-29).

Now the shaping filter problem, as considered by Darlington, is that of finding a weighting function $W_F(t,\tau)$ such that the systems corresponding to both it and its inverse are physically realizable and behave suitably as $\tau \to \infty$ for all t and such that

$$\Gamma_F(t_1, t_2) = \int_{-\infty}^{\infty} W_F(t_1,\tau)\ W_F^a(\tau, t_2)d\tau \qquad (1\text{-}30)$$

To do this he first finds the *bf's* of $B_F^{\Gamma}(p,t)$ and $A_F^{\Gamma}(p,t)$ from the *known bf's* of $B_S^{\Gamma}(p,t)$ and $B_N^{\Gamma}(p,t)$ and by solving (1-29). The problem then is to assign one-half of them to $W_F(t,\tau)$ and the remaining half to $W_F^a(t,\tau)$ so that the requirements demanded of $W_F(t,\tau)$ as stated above are met, if possible. Darlington shows that this is possible and shows how to do it providing the coefficients of the differential equations characterizing the systems used to generate S and N are regular at $t = \infty$, are periodic, or are of moderate variation. In these cases the bf's either become exponentials as $t \to \pm \infty$, are exponentials multiplied by periodic coefficients, or are dominated by exponentials as $t \to \pm \infty$ and those bf's associated with exponentials $e^{S\sigma t}$ where $R_e S_\sigma < 0$ are assigned to $W_p(t,\tau)$ just as in the stationary case, the $W_p(t,\tau)$ thereby obtained having the required properties.

Before proceeding to a summary of Batkov's work, it should be noted that in Darlington's work it was assumed that $\underline{T = (-\infty,\infty)}$ and that $\Gamma_p(t_1,t_2)$ was known to be the sum of two processes which were generated from uncorrelated *white noise* sources by *physically realizable* systems. Further, no terms due to initial conditions are present in $\Gamma_S(t_1,t_2)$ or $\Gamma_N(t_1,t_2)$ because of the assumption of stability of the S and N shaping filters and the choice of interval T.

1.2.3 Summary of the Work of Batkov*

In his paper [10], Batkov, like Darlington but from a somewhat different point of view, also studies the properties of weighting functions for physically realizable linear systems characterized by finite-order, linear differential equations and the properties of the covariance functions of the stochastic processes at their outputs when their inputs are *white noise* processes with the systems starting from rest; i.e., the initial conditions are zero. On the basis of this study, he presents three methods for solving the shaping filter problem for this class of nonstationary stochastic processes. The first of these methods is, in essence, that of Darlington but is not quite as fully developed as Darlington's. The second is an algebraic method using the discontinuities of the partial derivatives of $\Gamma(t_1,t_2)$ with respect to t_1 along the line $t_1 = t_2$; but, as mentioned earlier, it is much more restricted in application than claimed. The third uses a method due to Levy [15] for solving a certain type of nonlinear Volterra integral equation of the second kind by resolving kernels, the resolving kernals being solutions of linear Fredholm integral equations of the second kind. These methods and the results leading up to them are briefly summarized below.

Specifically, Batkov considers a physically realizable system characterized by the differential equation**

$$L(p,t)\,X(t) = M(p,t)\,Y(t) \tag{1-31}$$

*Batkov's paper contains several functions and operators with arguments incorrect.

**Many of the results stated above are derived in detail in Chapters 2 and 3 of this dissertation.

where

$$\left.\begin{aligned} L(p,t) &= \sum_{i=0}^{n} a_i(t)\,\frac{d^i}{dt^i}\,, \quad p = \frac{d}{dt}\,, \quad m < n \\ M(p,t) &= \sum_{j=0}^{m} b_j(t)\,\frac{d^j}{dt^j} \end{aligned}\right\} \tag{1-32}$$

and the $a_i(t)$ and $b_j(t)$ have n-derivatives in the region of interest. The weighting function $G(t,\tau)$ of a system characterized by (1-31) with $M(p,t) = 1$ is the solution of

$$L(p,t)\,G(t,\tau) = \delta(t-\tau) \tag{1-33}$$

with zero initial conditions. The weighting function, $G^a(t,\tau)$ of the corresponding adjoint system is the solution of

$$L^a(p,t)\,G^a(t,\tau) = -\,\delta(t-\tau)^* \tag{1-34}$$

with zero initial conditions and $G^a(t,\tau) = G(\tau,t)$. When $M(p,t)$ is not a constant, the weighting function $W(t,\tau)$ of the system is the solution of

$$L(p,t)\,W(t,\tau) = M(p,t)\,\delta(t-\tau) \tag{1-35}$$

with zero initial conditions and can be obtained from $G(t,\tau)$ by the expression

$$W(t,\tau) = M^a(p,\tau)\,G(t,\tau) \tag{1-36}$$

The weighting function $W^a(t,\tau)$ of the corresponding adjoint system is the solution of

$$L^a(p,t)\,W^a(t,\tau) = -\,M^a(p,t)\,\delta(t-\tau) \tag{1-37}$$

with zero initial conditions and $W_a(t,\tau) = W(\tau,t)$. As a function of τ, $W(t,\tau)$ satisfies the equation,

$$R(p,\tau)\,W(t,\tau) = Q(p,\tau)\,\delta(t-\tau) \tag{1-38}$$

where

$$\left.\begin{aligned} R(p,\tau) &= \sum_{i=0}^{n} r_i(\tau)\,\frac{d^i}{d\tau^i} \\ Q(p,\tau) &= \sum_{j=0}^{M} q_j(\tau)\,\frac{d^j}{d\tau^j} \end{aligned}\right\} \tag{1-39}$$

---

*The superscript a denotes the adjoint differential operator.

the coefficients $r_i(\tau)$ and $q_j(\tau)$ being determined from the coefficients $a_i(\tau)$ and $b_j(\tau)$ algebraically. Now $W(t,\tau)$ also is the solution of

$$L(p,t)\; W(t,\tau) \;=\; 0 \tag{1-40}$$

with the initial conditions

$$\partial^i W(t,\tau)/\partial t^i \Big|_{t\,\downarrow\,\tau} \;=\; 0 \quad ; \quad i = 0, \ldots, n-m-2 \tag{1-41}$$

$$\partial^{n-j-1} W(t,\tau)/\partial t^{n-j-1}\Big|_{t\,\downarrow\,\tau} \;=\; 1/a_n(\tau)\Big[\sum_{k=j}^{m} (-1)^{k-j} \binom{k}{j} b_j^{(k-j)}(\tau) \quad ;$$

$$- \sum_{r=n-m-1}^{n-j-2} \frac{\partial^r W(t,\tau)}{\partial t^r}\Big|_{t\,\downarrow\,\tau} \sum_{s=0}^{n-j-r-1} (-1)^s \binom{j+s}{j} a^{(s)}_{s+j+r+1}(\tau)\Big] \;;$$

$$j \;=\; 0, \ldots, m.$$

Again, $W(t,\tau)$ is also the solution of

$$R(p,\tau)\; W(t,\tau) \;=\; 0 \tag{1-42}$$

with the initial (final!) conditions

$$\partial^j W(t,\tau)/\partial\tau^j\Big|_{\tau\,\uparrow\,t} \;=\; \sum_{k=n-1}^{m+j} \frac{\partial^k G(t,\tau)}{\partial\tau^k}\Big|_{\tau\,\uparrow\,t} \;\times$$

$$\times \sum_{\substack{\ell=k-j\\ \ell>0}}^{m} (-1)^{\ell} \binom{\ell+j}{k} b_{\ell}^{(j+\ell-k)}(t), \; j \;=\; n-m-1, \ldots, n-1 \tag{1-43}$$

where

$$\partial^{n+k} G(t,\tau)/\partial\tau^{n+k}\Big|_{\tau\,\uparrow\,t} \;=\; 1/a_n(t) \sum_{\ell=n-1}^{m+k-1} \frac{\partial^{\ell} G(t,\tau)}{\partial\tau^{\ell}}\Big|_{\tau\,\uparrow\,t} \;\times$$

$$\times \sum_{\substack{i=\ell-k\\ i>0}}^{n} (-1)^{n-i-1} \binom{i+k}{\ell} a_i^{(k+i-1)}(t) \;; \quad k \;=\; 0, \ldots, m, \ldots \tag{1-44}$$

Moreover, $W(t,\tau)$ can be expressed in the form

$$W(t,\tau) = \begin{cases} \sum_{i=1}^{n} q_i(t)\ \beta_i(\tau) & ;\ t > \tau \\ 0 & ;\ t < \tau \end{cases} \tag{1-45}$$

where the $q_i(t)$ and $\beta_i(\tau)$ are sets of basis functions for $L(p,t)$ and $R(p,\tau)$ respectively.

Assuming $\{Y(t)\}$ is a *white noise* process applied to the system at time t_0 and that the system is at rest at t_0, the covariance function $\Gamma(t_1,t_2)$ of the output process $\{X(t)\}$ can be expressed in the form

$$\Gamma(t_1,t_2) = \begin{cases} \int_{t_0}^{t_2} W(t_1,\tau)\ W(t_2,\tau)\ d\tau & ;\ t_1 > t_2 \\ \int_{t_0}^{t_1} W(t_1,\tau)\ W(t_2,\tau)\ d\tau & ;\ t_1 < t_2 \end{cases} \tag{1-46}$$

Applying the operator $L(p,t_1)$ to (1-46) yields

$$L(p,t_1)\ \Gamma(t_1,t_2) = 0 \qquad ;\ t_1 > t_2 \tag{1-47}$$

$$L(p,t_1)\ \Gamma(t_1,t_2) = M(p,t_1)\ W(t_2,t_1) \qquad ;\ t_1 < t_2 \tag{1-48}$$

Further, $\Gamma(t_1,t_2)$ has $2n-2m-2$ continuous partial derivatives with respect to t_1 and t_2 and for $k \geq 2n-2m-1$

$$\Gamma_k(t_2,t_2) = \frac{\partial^k\Gamma(t_1,t_2)}{\partial t_1^k}\Big|_{t_1 \uparrow t_2} - \frac{\partial^k\Gamma(t_1,t_2)}{\partial t_1^k}\Big|_{t_1 \downarrow t_2} =$$

$$= \sum_{i=n-m-1}^{k-n+m} \ \sum_{\ell=n-m-1}^{i} \left\{ \binom{i}{\ell} \frac{\partial^{\ell} W(t_2,t_1)}{\partial t_1^{\ell}} \frac{\partial^{i-\ell}}{\partial t_1^{\,i-\ell}} \left[\frac{\partial^{k-i-1} W(t_1,\lambda)}{\partial t_1^{k-i-1}} \right]_{\lambda \uparrow t_1} \right\}_{t_1 \downarrow t_2} \tag{1-49}$$

Also, from (1-47) and (1-49), or from (1-45), it follows that

$$\Gamma(t_1,t_2) = \begin{cases} \sum_{i=1}^{n} q_i(t_1)\ p_i(t_2) & ;\ t_1 > t_2 \\ \sum_{i=1}^{n} q_i(t_2)\ p_i(t_1) & ;\ t_1 < t_2 \end{cases} \tag{1-50}$$

where the $q_i(t_1)$ are a set of basis functions for $L(p,t_1)$ and the $p_i(t_1)$ are a set of particular solutions of (1-48).

The first method suggested by Batkov for solving the shaping filter problem is based on (1-48). Given the $q_i(t_1)$, the $a_i(t_1)$ and $G(t,\tau)$ are easily computed algebraically (see Chapter 2). Knowing $L(p,t_1)$ and $G(t_2,t_1)$ and using (1-36), (1-48) becomes

$$L(p,t_1)\,\Gamma(t_1,t_2) \;=\; M(p,t_1)\,M^a(p,t_1)\,G(t_2,t_1);\;\; t_1 < t_2 \tag{1-51}$$

The remaining step is to find the product operator $M(p,t_1)\,M^a(p,t_1)$ from (1-51) and then decompose it into its adjoint factors. As Batkov points out, this decomposition is difficult if $M(p,t)$ contains derivative operators, but is simple if it is just a time-varying scale factor. Note the similarity of factoring the product $M(p,t_1)\,M^a(p,t_1)$ here and the factoring of $A_p^{\Gamma}(p,t)$ in Darlington's work. These are, in essence, the same problem.

The second method described by Batkov, i.e., the algebraic method, is based on (1-49) rewritten in the form

$$\begin{aligned}\frac{\partial^j W(t,\tau)}{\partial\tau^j} = {} & \frac{a_n(t)}{b_m(t)}\left\{\Gamma_{j+n-m}(t,t) - \sum_{i=n-m-1}^{j-1}\left[\sum_{\ell=n-m-1}^{i}\binom{i}{\ell}\times\right.\right.\\ & \frac{\partial^\ell W(t,\tau)}{\partial\tau^\ell}\;\frac{d^{i-\ell}}{dt^{i-\ell}}\left(\frac{\partial^{j+n-m-i-1}W(t,\lambda)}{\partial t^{j+n-m-i-1}}\Big|_{\lambda=t}\right) + \binom{j}{i}\frac{\partial^i W(t,\tau)}{\partial\tau^i}\times\\ & \left.\left.\frac{d^{j-i}}{d^{j-i}}\;\frac{b_m(t)}{b_n(t)}\right]_{\tau=t}\right\}\;;\;\; j = n-m-1,\,.\,.\,.\,,n-1\end{aligned} \tag{1-52}$$

From (1-52), relationships between the partial derivatives of $W(t,\tau)$ with respect to t and τ for $t=\tau$ can be found recursively and expressing them in terms of the known $a_i(t)$ and unknown $b_\ell(t)$ lead successively to m-equations in the $b_\ell(t)$ and their derivatives. Batkov claims that $b_{m-k}(t)$ enters the equation obtained from (1-52) for $j = n-m+k-1$ algebraically in terms of $b_m(t)$, . . . , $b_{m-k+1}(t)$ and their derivatives. Now, in particular,

$$b_m(t) \;=\; \pm\, a_n(t)\sqrt{(-1)^{n-m-1}\,\Gamma_{2n-2m-1}(t,t)} \tag{1-53}$$

However, as will be seen later, $b_{m-1}(t)$ does not appear in (1-52) for $j=n-m$ and both $b_{m-1}(t)$ and $b_{m-1}^{(1)}(t)$ and also $b_{m-2}(t)$ appear in (1-52) for $j=n-m+1$, etc., for $j=n-m+1$, . . . , $n-1$. Hence, the claimed recursive algebraic method for finding the $b_{m-k}(t)$ from (1-52), and thereby solving the shaping filter problem algebraically, fails (except, of course, where $b_j(t) \equiv 0$ for $j \neq 0$).

The third method, described by Batkov in an appendix to his paper, will not be summarized at this point because of the difficulty of solving the Fredholm integral equation for the resolvent kernel and because of his *restrictive* assumptions as noted in the following remarks.

Before proceeding, it should be noted that Batkov assumed that the *system started from rest at* t_0 where $T = (t_0,\infty)$ and, what is even more important, he also assumed that $\Gamma(t_1,t_2)$ is the covariance function of a process generated from a *white noise* process by a physically realizable linear system characterized by a finite-order linear differential equation of the form (1-31) for $t \geq t_0$.

1.2.4 Summary of the Work of Leonov

As noted in the chronological sketch, Leonov [12] has obtained a rather nice mathematical solution to the shaping filter problem, and also to the corresponding inverse shaping filter problem, in terms of expansions in orthogonal functions. This work is summarized here and the details can, of course, be found in Leonov's paper.

Leonov formulates the shaping filter problem as follows. Given a *white noise* process $\{Y(t)\}$* where $-\infty < t < \infty$ [i.e., $T_Y = (-\infty,\infty)$] and a nonstationary process $\{X(t)\}$ where $0 < t < T$ [i.e., $T_X = (0,T)$], it is required to show that the random function [a sample function of $\{X(t)\}$] $X(t)$ can, under certain conditions, be represented in the form

$$X(t) = A_X\, Y(t) \tag{1-54}$$

where the (linear) operator A_X is defined if the function $X(t)$ is given.** The corresponding inverse problem is that of representing $Y(t)$ in the form

$$Y(t) = A_X^{-1}\, X(t) \tag{1-55}$$

where A_X^{-1} is the operator inverse to A_X. Leonov shows that this can be done by explicitly constructing a suitable A_X and a suitable A_X^{-1} as follows.

As is well known [16] a random function $Z(t)$, $T_Z = (a,b)$, can be represented as a series (canonical expansion)

$$Z(t) = \sum_{i=0}^{\infty} B_i\, z_i(t)^{***} \tag{1-56}$$

---

*It is always assumed that $E\,Y(t) = 0$ for all *white noise* processes considered herein.

**Pugachev [16,17] calls (1-54) the integral canonical representation of $X(t)$.

***The $z_i(t)$ are not necessarily orthogonal and $-\infty \leq a < b \leq \infty$.

where the B_i are random variables which satisfy the conditions

$$E\, B_i\, B_j \;=\; \delta_{ij}\, D_j \tag{1-57}$$

and the $z_i(t)$ are some regular (nonrandom) functions. In order that the series in (1-56) converge in the mean to $Z(t)$, it is necessary and sufficient that the series

$$\Gamma_z(t_1,t_2) \;=\; \sum_{i=1}^{\infty} D_i\, z_i(t_1)\, z_j(t_2) \tag{1-58}$$

converge to $\Gamma_z(t_1,t_2)$ in the usual sense. The definition of convergence in the mean is, of course, only meaningful for random functions with finite variances.

Now to solve the problem, it is necessary to represent $Y(t)$ by a series of the form in (1-56). However, since $Y(t)$ does not have a finite variance, convergence in the mean cannot be used and a new concept of convergence must be introduced. Leonov introduces the concept of weak convergence in the mean.* A sequence of random functions $U_n(t)$ is said to converge weakly in the mean to the random function $U(t)$ if the integral

$$a_n(T) \;=\; \int_0^T R(t)\, U_n(t)\, dt \tag{1-59}$$

has a limit in the mean square sense as $n \to \infty$ for any *sufficiently smooth* random function $R(t)$; i.e., for any $R(t)$ which has finite variance, is continuous in the mean, has the necessary number of continuous stochastic derivatives, and whose covariance function $\Gamma_R(t_1,t_2)$ satisfies the inequality $\int_{-\infty}^{\infty} \Gamma_R^2(t,t)\, dt < \infty$. With this definition of convergence, Leonov shows that $Y(t)$ can be represented in the form

$$Y(t) \;=\; \sum_{i=1}^{\infty} C_i\, y_i(t) \tag{1-60}$$

where $E\, C_i C_j = \delta_{ij}$ and the $y_i(t)$ are any complete (in L_2) set of orthonormal functions over $(-\infty,\infty)$ and where the series in (1-60) converges weakly in the mean to the *white noise* random function $Y(t)$.

To solve the basic problem is now fairly easy. The C_i in (1-60) are defined as follows

$$C_i \;=\; \frac{V_i}{\sqrt{D_i}} \tag{1-61}$$

---

*This is clearly analogous to the ordinary concept of weak convergence in Hilbert space [18].

where the random variables are the coefficients in the series expansion of $X(t)$

$$X(t) = \sum_{i=1}^{\infty} V_i \, X_i(t) \tag{1-62}$$

and $D_i = E\, V_i^2$. The linear operator A_X is then defined as

$$A_X \, Y(t) = \int_{-\infty}^{\infty} W_X(t,\tau) \, Y(\tau) \, d\tau \tag{1-63}$$

where

$$W_X(t,\tau) = \sum_{i=1}^{\infty} \sqrt{D_i} \quad X_i(t) \, y_i(\tau) \tag{1-64}$$

Then from (1-63) and (1-64)

$$A_X \, Y(t) = \int_{-\infty}^{\infty} W_X(t,\tau) \, Y(\tau) \, d\tau = \sum_{i=1}^{\infty} V_i \, X_i(t) = X(t) \tag{1-65}$$

where the integral in (1-63) is taken in the mean.

The functions $W_X(t,\tau)$ and $Y(\tau)$ in (1-65) can be defined in infinitely many ways by using any other representation of $X(t)$ in the form (1-62) as is shown to be possible in [16]. However, Leonov shows that if $Y(t)$ is so chosen that (1-65) holds, then there is one and only one $W_X(t,\tau)$; i e., $W_X(t,\tau)$ is unique.

Finally, the inverse problem is easily solved as follows. Let A_X^{-1} be defined as

$$A_X^{-1} \, X(t) = \int_{-\infty}^{\infty} W_X^{-1}(t,\tau) \, X(\tau) \, d\tau \tag{1-66}$$

where

$$W_X^{-1}(t,\tau) = \sum_{i=1}^{\infty} y_i(t) \, a_i(\tau)/\sqrt{D_i} \tag{1-67}$$

and the $a_i(\tau)$ are chosen so that

$$\int_0^T a_i(\tau) \, X_j(\tau) \, d\tau = \delta_{ij} \tag{1-68}$$

As before, the $y_i(t)$ are any complete (in L_2) set of orthonormal functions over $(-\infty,\infty)$. Then from (1-60), (1-66), (1-67), and (1-68) it follows that

$$Y(t) = \int_0^T W_X^{-1}(t,\tau) \, X(\tau) \, d\tau = \sum_{i=1}^{\infty} \frac{V_i}{\sqrt{D_i}} \, y_i(t) \tag{1-69}$$

The series in (1-69) converges weakly in the mean to the *white noise* random function $Y(t)$ as noted above.

This completes the summary of Leonov's solution to the shaping filter problem and the corresponding inverse shaping filter problem, the remainder of his paper being devoted to mathematical niceties and applications.* It is clear that the weighting function for the shaping filter can be written down immediately in series form once the $X_i(t)$ and D_i for the canonical expansion of $X(t)$ are known. In his book [16] Pugachev presents several techniques for finding the first n-terms of expansions of the form (1-62) rather simply and which avoid having to determine the eigenvalues and eigenfunctions of an integral equation as required in the well known Karhunen-Loeve Expansion Theorem. However, it should be noted that Leonov's solution is always obtained in the form of an infinite series and, further, there is no guarantee of physical realizability of the shaping filter or its inverse.

*Applications will be discussed in later sections.

CHAPTER 2

PROPERTIES OF WEIGHTING FUNCTIONS FOR A CLASS OF LINEAR SYSTEMS

2.1 INTRODUCTION

Because of their importance in the shaping filter as evidenced by Chapter 1 and those to follow, for purposes of detailed review and for later use as available reference material, this chapter is devoted to the investigation of the properties of weighting functions (Green's functions) for systems which can be described by finite-order ordinary linear differential equations of the form

$$\sum_{i=0}^{n} a_i(t)\, X^{(i)}(t) \;=\; \sum_{j=0}^{n-1} b_j(t)\, y^{(j)}(t) \quad ; \; t > 0 \tag{2-1}$$

where, for $t > 0$, $a_n(t) \neq 0$ and the $a_i(t)$ and $b_j(t)$ are continuous. Since weighting functions and their derivatives are, in general, discontinuous at $t \neq \tau$, this investigation divides naturally into two parts: properties in the regions where $t \neq \tau$, and properties of the discontinuities at $t = \tau$.

2.2 PROPERTIES IN THE REGIONS WHERE $t \neq \tau$

It is well known that the general solution of (2-1) can be written in the form [19, p.257]

$$X(t) \;=\; \int_0^t d\tau\, G(t,\tau) \sum_{j=0}^{n-1} b_j(\tau)\, y^{(j)}(\tau) \;+ \\ + \sum_{i=0}^{n-1} X^{(i)}(0)\, q_i(t) \tag{2-2}$$

where the $q_i(t)$ are n-independent solutions of (2-3)

$$\sum_{i=0}^{n} a_i\, X^{(i)}(t) \;=\; 0 \tag{2-3}$$

for which

$$q_i^{(j)}(t)\Big|_{t \downarrow 0} \;=\; \delta_{ij}; \; i,j \;=\; 0, \ldots, n-1 \tag{2-4}$$

and $G(t,\tau)$ is the weighting function (Green's function) for (2-3). If $G(t,\tau)$

and the $b_j(\tau)$ have a sufficient number of derivatives with respect to τ, then through integration by parts, [20, p.189] (2-2) can be brought into the form

$$X(t) = \int_0^t d\tau\, W(t,\tau)\, y(\tau) + \sum_{i=0}^{n-1} X^{(i)}(0) q_i(t) \tag{2-5}$$

where

$$W(t,\tau) = \sum_{j=0}^{n-1} (-1)^j \frac{\partial^j b_j(\tau)\, G(t,\tau)}{\partial \tau^j} \tag{2-6}$$

Here, $W(t,\tau)$ is the weighting function for (2-1). Its properties are investigated below.

Since $W(t,\tau)$ is defined in terms of $G(t,\tau)$ as shown in (2-6), any investigation of the properties of $W(t,\tau)$ must begin with an investigation of the properties of $G(t,\tau)$. By definition [19, p.254], the weighting function for (2-3) is that solution $G(t,\tau)$ of (2-3) which satisfies the condition

$$\begin{aligned} G(t,\tau) &= 0 \quad ; \quad t < \tau \\ \lim_{t \downarrow \tau} \frac{\partial^i G(t,\tau)}{\partial t^i} &= 0 \quad ; \quad i = 0, \ldots, n-2 \\ \lim_{t \downarrow \tau} \frac{\partial^{n-1} G(t,\tau)}{\partial t^{n-1}} &= \frac{1}{a_n(\tau)} \end{aligned} \tag{2-7}$$

Since $G(t,\tau)$ is a solution of (2-3) for $t > \tau$, in this region it can, according to the theory of linear differential equations, be expressed as a linear combination of the $q_i(t)$.* Hence

$$G(t,\tau) = \sum_{i=0}^{n-1} \alpha_i\, q_i(t) \quad ; \quad t > \tau \tag{2-8}$$

where the α_i are chosen so as to satisfy conditions (2-7). Putting (2-8) into (2-7) and writing the result in vector-matrix form there results

$$\begin{bmatrix} q_0(\tau) & \ldots & q_{n-1}(\tau) \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ q_0^{(n-1)}(\tau) & \ldots & q_{n-1}^{(n-1)}(\tau) \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \cdot \\ \cdot \\ \cdot \\ \alpha_{n-1} \end{bmatrix} = \begin{bmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ 1/a_n(\tau) \end{bmatrix} \tag{2-9}$$

*Or any other set of n-linearly independent solutions of (2-3).

Now the determinant of the matrix in (2-9) is the Wronskian of the $q_i(t)$ and hence does not vanish because the $q_i(t)$ are a set of *linearly independent* solutions of (2-3). Therefore, the inverse of the matrix in (2-9) exists and, hence, the a_i are given by

$$\begin{bmatrix} a_0 \\ \cdot \\ \cdot \\ \cdot \\ a_{n-1} \end{bmatrix} = \begin{bmatrix} q_0(\tau) & \cdots & q_{n-1}(\tau) \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ q_0^{(n-1)}(\tau) & \cdots & q_{n-1}^{(n-1)}(\tau) \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ 1/a_n(\tau) \end{bmatrix} \tag{2-10}$$

It is seen from (2-10) that the a_i depend on τ only, which, together with (2-8), implies that $G(t,\tau)$ is *separable*; i.e., that

$$G(t,\tau) = \begin{cases} \sum_{i=0}^{n-1} a_i(\tau)\, q_i(t) & ; \quad t > \tau \\ 0 & ; \quad t < \tau \end{cases} \tag{2-11}$$

Under the assumptions made up to now regarding the $a_i(t)$, it follows from the existence theorem for solutions of a differential equation that the $q_i(t)$ have at least n-continuous derivatives for $t > 0$, and, thus, from (2-10) it follows the $a_i(\tau)$ are at least continuous for $\tau > 0$. More generally, if the $a_i^{(i)}(t)$ are continuous $t > 0$, then the $a_i(\tau)$ have at least n-continuous derivatives for $\tau > 0$. To show this, it is convenient to introduce the adjoint differential equation corresponding to (2-3); namely,

$$\sum_{i=0}^{n} (-1)^i\, [a_i(t)\, X(t)]^{(i)} = 0 \tag{2-12}$$

Since the $a_i^{(i)}(t)$ are assumed to be continuous, (2-12) can be rewritten in the form

$$\sum_{i=0}^{n} C_i(t)\, X^{(i)}(t) = 0 \tag{2-13}$$

where the $C_i(t)$ are continuous for $t > 0$. Now, if $H(t,\tau)$ is used to denote the weighting function for (2-13), then [19, p.256]

$$H(t,\tau) = G(\tau,t) \tag{2-14}$$

Furthermore, $H(t,\tau)$ must satisfy (2-13) for all $t > \tau$. Hence, using (2-11), it is found that

$$\sum_{i=0}^{n-1} q_i(\tau) \sum_{j=0}^{n} C_j(t)\, a_i^{(j)}(t) = 0 \quad ; \quad t > \tau \tag{2-15}$$

Since the $q_i(\tau)$ are linearly independent, (2-15) implies that

$$\sum_{j=0}^{n} C_j(t)\, a_i^{(j)}(t) = 0 \tag{2-16}$$

Thus, the $a_i(t)$ are solutions of (2-13) and, hence, have n-continuous derivatives by the existence theorem.

From (2-16), it is easily shown that

$$a_i^{(n)}(t) = -\frac{\sum_{j=0}^{n-1} C_j(t)\, a_i^{(j)}(t)}{C_n(t)} \tag{2-17}$$

From (2-17), it is easily concluded that, if the $C_i^{(k)}(t)$ are continuous [i.e., the $a_i^{(i+k)}(t)$ are continuous], then the $a_i^{(n+k)}(t)$ are continuous. Similarly, from (2-3) it can be concluded that the $q_i^{(n+k)}(t)$ are continuous if the $a_i^{(k)}(t)$ are continuous. This concludes the study of $G(t,\tau)$.

Substituting (2-11) into (2-6) it is easily seen that $W(t,\tau)$ can be written in the form

$$W(t,\tau) = \begin{cases} \sum_{i=0}^{n} q_i(t) \sum_{j=0}^{n-1} (-1)^j [b_j(\tau) a_i(\tau)]^{(j)} & ; \ t > \tau \\ 0 & ; \ t < \tau \end{cases} \tag{2-18}$$

The existence of the derivatives in (2-18) will be guaranteed by assuming continuity of the $a_i^{(i)}(t)$ and $b_j^{(j)}(t)$ for $t > 0$. Carrying out the indicated differentiation and m'king the obvious definitions for the $\beta_i(\tau)$, it is found that $W(t,\tau)$ can be written in the form

$$W(t,\tau) = \begin{cases} \sum_{i=0}^{n-1} q_i(t)\beta_i(\tau) & ; \ t > \tau \\ 0 & ; \ t < \tau \end{cases} \tag{2-19}$$

Hence, $W(t,\tau)$ is *separable;* i.e., can be written as the sum of products of a function of t only by a function of τ only. Further, the functions of t are a set of linearly independent solutions of (2-3). It is clear that continuity of the derivatives of the $q_i(t)$ and $\beta_i(\tau)$ implies the continuity of the partial derivatives of $W(t,\tau)$ at any t,τ except $t = \tau$. Examining the differentiability properties of the $\beta_i(\tau)$ it is found that the $\beta_i^{(k)}(\tau)$ are continuous for $\tau > 0$

if the $a_i^{(i+k)}(t)$ and $b_j^{(j+k)}(t)$ are continuous for $t > 0$. Incidentally, it is also clear that $W(t,\tau)$ formally satisfies (2-3) except at $t = \tau$.

2.3 PROPERTIES OF THE DISCONTINUITIES OF $W(t,\tau)$ AND ITS PARTIAL DERIVATIVES AT $t = \tau$

In order to investigate the discontinuities of $W(t,\tau)$ and its partial derivatives at $t = \tau$, it will be convenient to replace (2-1) by a particular set of n first-order equations. If the $a_i(t)$ and $b_j(t)$ have a sufficient number of continuous derivatives for $t > 0$, then (2-1) is equivalent to the following system of first-order equations (see [20, p.191]).

$$
\begin{aligned}
X(t) &= X_1(t) + F_0(t)\, y(t) \\
X_1^{(1)}(t) &= X_2(t) + F_1(t)\, y(t) \\
&\;\;\vdots \\
X_{n-1}^{(1)}(t) &= X_n(t) + F_{n-1}(t)\, y(t) \\
X_n^{(1)}(t) &= -a_{n-1}(t)\, X_n(t) \ldots - a_0(t)\, X_1(t) + F_n(t)\, y(t)
\end{aligned}
\tag{2-20}
$$

where the $F_i(t)$ can be found recursively from the equation

$$
F_i(t) = b_{n-i}(t) - \sum_{k=1}^{i-1} \sum_{S=0}^{i-k} \binom{n+S-i}{n-i} a_{n-i+k+S}(t)\, F_k^{(S)}(t) \tag{2-21}
$$

In (2-21) use has been made of the fact that $F_0(t) = b_n(t) \equiv 0$ and, hence, all terms involving $F_0(t)$ have been omitted. If $W_i(t,\tau)$ is used to denote the weighting function corresponding to $X_i(t)$ in (2-20), then letting $y(t) = \delta(t-\tau)$, (2-20) yields

$$
\begin{aligned}
W(t,\tau) &= W_1(t,\tau) \\
W_1^{(1,0)}(t,\tau) &= W_2(t,\tau) + F_1(t)\, \delta(t-\tau) \\
&\;\;\vdots \\
W_{n-1}^{(1,0)}(t,\tau) &= W_n(t,\tau) + F_{n-1}(t)\, \delta(t-\tau) \\
W_n^{(1,0)}(t,\tau) &= -\sum_{i=1}^{n} a_{i-1}(t)\, W_i(t,\tau) + F_n(t)\, \delta(t-\tau)
\end{aligned}
\tag{2-22}
$$

Integrating (2-22) with respect to t from 0 to t and remembering that $W_i(t,\tau)=0$

for $t < \tau$, there results (for $\tau > 0$)

$$\begin{aligned}
\int_{\tau}^{t} d\sigma\, W(\sigma,\tau) &= \int_{\tau}^{t} d\sigma\, W_1(\sigma,\tau) \\
W_1(t,\tau) &= \int_{\tau}^{t} d\sigma\, W_2(\sigma,\tau) + F_1(\tau)\, U(t-\tau) \\
&\vdots \\
W_{n-1}(t,\tau) &= \int_{\tau}^{t} d\sigma\, W_n(\sigma,\tau) + F_{n-1}(\tau)\, U(t-\tau) \\
W_n(t,\tau) &= \int_{\tau}^{t} d\sigma \left[-\sum_{i=1}^{n} a_{i-1}(\sigma)\, W_i(\sigma,\tau) \right] + F_n(\tau) U(t-\tau)
\end{aligned} \qquad \begin{matrix}(2\text{-}23)\\ \\ (2\text{-}24)\end{matrix}$$

where $U(t-\tau)$ is the ordinary unit step function. Making use of (2-22) and (2-23), the discontinuities of $W(t,\tau)$ and its partial derivatives can be evaluated fairly easily. In fact from (2-22) it is easily established that

$$\begin{aligned}
W^{(i,0)}(t,\tau) &= W_{i+1}(t,\tau) \;:\; t > \tau,\; i = 0, \ldots, n-1 \\
W^{(n,0)}(t,\tau) &= -\sum_{i=1}^{n} a_{i-1}(t)\, W_i(t,\tau) \;:\; t > \tau
\end{aligned} \qquad (2\text{-}24)$$

If $W^{(i,j)}(\tau,\tau)$ is defined as shown in (2-25)

$$W^{(i,j)}(\tau,\tau) = \lim_{t \downarrow \tau} \frac{\partial^{i+j} W(t,\tau)}{\partial t^i\, \partial \tau^j} \qquad (2\text{-}25)$$

then from (2-23) and (2-24) it is clear that

$$\begin{aligned}
W^{(i,0)}(\tau,\tau) &= F_{i+1}(\tau) \qquad ; \quad i = 0, \ldots, n-1 \\
W^{(n,0)}(\tau,\tau) &= -\sum_{i=1}^{n} a_{i-1}(\tau)\, F_i(\tau)
\end{aligned} \qquad (2\text{-}26)$$

Also from (2-24) it is clear that

$$\begin{aligned}
W^{(i,j)}(t,\tau) &= W_{i+1}^{(0,j)}(t,\tau) \;;\; t > \tau \;;\; i = 1, \ldots, n-1 \qquad (2\text{-}27) \\
W^{(n,j)}(t,\tau) &= -\sum_{i=1}^{n} a_{i-1}(t)\, W_i^{(0,j)}(t,\tau) \;;\; t > \tau
\end{aligned}$$

By virtue of (2-27), the evaluation of the discontinuities of the $W^{(i,j)}(t,\tau)$ at $t = \tau$ reduces to the evaluation of the discontinuities of the $W_i^{(0,j)}(t,\tau)$ at $t = \tau$. To find these, use is made of (2-23). Upon performing the indicated differentiation, it is found that

$$W_i^{(0,j)}(t,\tau) = \int_\tau^t d\sigma\, W_{i+1}^{(0,j)}(\sigma,\tau) - \sum_{k=0}^{j-1} \left[W_{i-1}^{(0,k)}(\tau,\tau) \right]^{(j-1-k)}$$

$$+ \quad F_i^{(j)}(\tau) \ ; \ t > \tau, \ i < n$$

$$W_n^{(0,j)}(t,\tau) = \int_\tau^t d\sigma \left[- \sum_{i=1}^{n} a_{i-1}(\sigma)\, W_i^{(0,j)}(\sigma,\tau) \right]$$

$$- \sum_{k=0}^{j-1} \left[- \sum_{i=1}^{n} a_{n-1}(\tau)\, W_i^{(0,k)}(\tau,\tau) \right]^{(j-1-k)} \tag{2-28}$$

$$+ \quad F_n^{(j)}(\tau) \ ; \ t > \tau$$

Taking the limit as $t \downarrow \tau$, there results

$$W_i^{(0,k)}(\tau,\tau) = - \sum_{k=0}^{j-1} \left[W_{i+1}^{(0,k)}(\tau,\tau) \right]^{(j-1-k)} + F_i^{(j)}(\tau) \ ; \ i < n$$

$$W_n^{(0,j)}(t,\tau) = - \sum_{k=0}^{j-1} \left[- \sum_{i=1}^{n} a_{i-1}(\tau)\, W_i^{(0,k)}(t,\tau) \right]^{(j-1-k)}$$

$$+ \quad F_n^{(j)}(\tau) \tag{2-29}$$

Equations (2-29) can be used to find the values of $W_i^{(0,j)}(\tau,\tau)$, and, hence, those of $W^{(i,j)}(\tau,\tau)$, recursively in terms of the $F_i(\tau)$ and $a_i(\tau)$ and their derivatives.

2.4 A SYNTHESIS PROBLEM

This chapter concludes with the solution of the following synthesis problem; Given $W(t,\tau)$ expressed in the form in (2-19), find the $a_i(t)$ and $b_j(t)$ of its corresponding differential equation (2-1).

Since the $q_i(t)$ in (2-19) are solutions of (2-3), it is clear that the following equations are valid.

$$\begin{aligned} \sum_{i=0}^{n} a_i(t)\, q_0^{(i)}(t) &= 0 \\ &\ \vdots \\ \sum_{i=0}^{n} a_i(t)\, q_{n-1}^{(i)}(t) &= 0 \end{aligned} \tag{2-30}$$

Regarding the $q_i(t)$ as known and the $a_i(t)$ as unknown, (2-30) is a set of n-linear equations in the $n + 1$ unknowns, $a_i(t)$. Since it can be assumed without loss of generality in (2-1) that $a_n(t) = 1$, (2-30) can be solved for the remaining unknown $a_i(t)$ with the result

$$\begin{bmatrix} a_0(t) \\ \cdot \\ \cdot \\ \cdot \\ a_{n-1}(t) \end{bmatrix} = \begin{bmatrix} q_0(t) & \cdots & q_0^{(n-1)}(t) \\ \cdot & & \\ \cdot & & \\ \cdot & & \\ q_{n-1}(t) & \cdots & q_{n-1}^{(n-1)}(t) \end{bmatrix}^{-1} \times \begin{bmatrix} -q_0^{(n)}(t) \\ \cdot \\ \cdot \\ \cdot \\ -q_{n-1}^{(n)}(t) \end{bmatrix} \quad (2\text{-}31)$$

The existence of the inverse matrix in (2-31) is guaranteed by the fact that the $q_i(t)$ are linearly independent solutions of (2-3). Equation (2-31) provides the desired relationship for determining the $a_i(t)$ from $W(t,\tau)$. This result is given as Theorem 6.2 in [21] for an arbitrary set of n-line rly independent functions $q_i(t)$ providing the $q_i(t)$ have n-continuous derivatives on the region of interest.* Once the $a_i(t)$ have been determined, the $b_j(t)$ can be found from (2-21) rewritten in the form

$$b_{n-i}(t) = \sum_{k=1}^{i} \sum_{S=0}^{i-k} \binom{n+S-i}{n-i} a_{n-i+k+S}(t)\, F_k^{(S)}(t) \quad (2\text{-}32)$$

where the $F_i(t)$ are given directly in terms of the discontinuities of the weighting function and its derivatives with respect to t as in (2-26).

A development similar to that given in Sections 2.2, 2.3, and 2.4 has also been carried out by Borskii [22].

*Linear independence of the $q_i(t)$ is equivalent to the nonvanishing of their Wronskian.

CHAPTER 3

PROPERTIES OF THE COVARIANCE FUNCTIONS OF A CLASS OF STOCHASTIC PROCESSES

3.1 INTRODUCTION

As is to be expected and as evidenced by Chapter 1, the covariance functions of the class of stochastic processes, which can be generated by passing *white noise* processes through systems characterizable by finite-order ordinary linear differential equations of the type given in (2-1) with random initial conditions, have many properties in common in addition to that of being non-negative-definite [23,p.466]. This chapter is devoted to the development of some of the more interesting and useful among these additional properties both for the purpose of detailed review and later use.

3.2 SOME GENERAL PROPERTIES

Let it be assumed that a given system can be characterized by a differential equation of the form given in (2-1) where $a_n(t) = 1$ and the $a_i{}^{(i)}(t)$ and $b_j{}^{(j)}(t)$ are continuous for $t > 0$. Then the weighting function, $W(t,\tau)$, for the systems exists in the form (2-19). Now, letting $\{y(t)\}$ be a stochastic process with covariance function $\Gamma_{yy}(t_1,t_2)$ and $X^{(i)}(0)$, random variables with covariances Γ_{ij}, the covariance function, $\Gamma_{XX}(t_1,t_2)$, of the stochastic process $\{X(t)\}$ can be determined from $\Gamma_{yy}(t_1,t_2)$ and Γ_{ij} according to the relation [20,p.227].*

$$\Gamma_{XX}(t_1,t_2) = \int_0^{t_1} d\tau_1 \int_0^{t_2} d\tau_2 W(t_1,\tau_1)\, W(t_2,\tau_2)\Gamma_{yy}(\tau_1,\tau_2) + \sum_{i=0}^{n-1}\sum_{j=0}^{n-1} \Gamma_{ij}\, q_i(t_1)\, q_j(t_2) \quad ; \; t_1 \geq 0,\; t_2 \geq 0 \tag{3-1}$$

When $y(t)$ is a *white noise* process, $\Gamma_{yy}(t_1,t_2) = \delta(t_1 - t_2)$ and (3-1) becomes

$$\Gamma_{XX}(t_1,t_2) = \begin{cases} \displaystyle\int_0^{t_2} d\tau\, W(t_1,\tau)\, W(t_2,\tau) + \sum_{i=0}^{n-1}\sum_{j=0}^{i-1} \Gamma_{ij}\, q_i(t_1)\, q_j(t_2)\; ; & t_1 > t_2 \geq 0 \\ \displaystyle\int_0^{t_1} d\tau\, W(t_1,\tau)\, W(t_2,\tau) + \sum_{i=0}^{n-1}\sum_{j=0}^{n-1} \Gamma_{ij}\, q_i(t_1)\, q\,(t_2); & t_2 > t_1 \geq 0 \end{cases} \tag{3-2}$$

---

*Here it has been assumed that $E[X^{(i)}(0)y(t)] = 0$ in i and $t \geq 0$.

If the expression for $W(t,\tau)$ given in (2-19) is substituted into (3-2) and the coefficients of the $q_i(t)$ are collected, (3-2) can be brought into the form

$$\Gamma_{XX}(t_1, t_2) = \begin{cases} \sum_{i=0}^{n-1} q_i(t_1) \sum_{j=0}^{n-1} q_j(t_2) \left[\int_0^{t_2} d\tau \beta_i(\tau)\beta_j(\tau) + \Gamma_{ij} \right] ; & t_1 > t_2 \geq 0 \\ \sum_{j=0}^{n-1} q_j(t_2) \sum_{i=0}^{n-1} q_i(t_1) \left[\int_0^{t_1} d\tau \beta_i(\tau)\beta_j(\tau) + \Gamma_{ij} \right] ; & t_2 > t_1 \geq 0 \end{cases} \tag{3-3}$$

It will be convenient to denote the terms in the square brackets in (3-3) by the symbols $p_i(t)$, in which case (3-3) becomes

$$\Gamma_{XX}(t_1, t_2) = \begin{cases} \sum_{i=0}^{n-1} q_i(t_1)\, p_i(t_2) & ; \quad t_1 > t_2 \geq 0 \\ \sum_{j=0}^{n-1} q_j(t_2)\, p_j(t_1) & ; \quad t_2 > t_1 \geq 0 \end{cases} \tag{3-4}$$

From (3-4) two important general properties of the class of covariance functions under consideration are obvious. First, they are separable in the sense used in the preceding section. Second, for $t_1 > t_2$, the functions of t_1, i.e., the $q_i(t_1)$, are solutions of the homogeneous differential equation (2-3). In addition, if the $a_i^{(i+k-1)}(t)$ and $b_j^{(j+k-1)}(t)$ are continuous for $t > 0$, then $\Gamma_{XX}(t_1, t_2)$ has continuous partial derivatives of order k in regions t_1, $t_2 > 0$; $t_1 \neq t_2$. This is easily established from the differentiability properties of the $q_i(t)$ and $\beta_j(\tau)$ discussed in Section 2.2.

3.3 DISCONTINUITY PROPERTIES

The discontinuities of the partial derivatives of $\Gamma_{XX}(t_1, t_2)$ at $t_1 = t_2$ can be evaluated in terms of the discontinuities of the partial derivatives of $W(t,\tau)$ at $t = \tau$ by making use of (3-2). In fact, upon differentiating (3-2) partially with respect to t_1 (assuming, of course, that the requisite derivatives exist) there results

$$\Gamma_{XX}^{(1,0)}(t_1,t_2) = \begin{cases} \int_0^{t_2} d\tau W^{(1,0)}(t_1,\tau)\, W(t_2,\tau) + \\ + \sum_{i=0}^{n-1}\sum_{j=0}^{n-1} \Gamma_{ij}\, q_i^{(1)}(t_1)\, q_j(t_2) \quad ; \quad t_1 > t_2 \\ \\ \int_0^{t_1} d\tau W^{(1,0)}(t_1,\tau)\, W(t_2,\tau) + W(t_1,t_1)\, W(t_2,t_1) + \\ + \sum_{i=0}^{n-1}\sum_{j=0}^{n-1} \Gamma_{ij}\, q_i^{(1)}(t_1)\, q_j(t_2) \;;\; t_2 > t_1 \end{cases} \tag{3-5}$$

Defining the discontinuities $J_{ij}(t_2)$ as shown in (3-6)

$$J_{i,j}(t_2) = \lim_{t_1 \downarrow t_2} \Gamma_{XX}^{(i,j)}(t_1,t_2) - \lim_{t_1 \uparrow t_2} \Gamma_{XX}^{(i,j)}(t_1,t_2) \tag{3-6}$$

it follows from (3-5) that

$$J_{1,0}(t_2) = W^2(t_2,t_2) \tag{3-7}$$

In general

$$\Gamma_{XX}^{(i,j)}(t_1,t_2) = \begin{cases} \int_0^{t_2} d\tau W^{(i,0)}(t_1,\tau)\, W^{(j,0)}(t_2,\tau) + \\ + \sum_{\ell=0}^{j-1} \frac{\partial^\ell}{\partial t_2^\ell}[W^{(i,0)}(t_1,t_2)\, W^{(j-\ell-1,0)}(t_2,t_1)] + \\ + \sum_{k=0}^{n-1}\sum_{\ell=0}^{n-1} \Gamma_{k\ell}\, q_k^{(i)}(t_1)\, q_\ell^{(j)}(t_2) \;;\; t_1 > t_2 \\ \\ \int_0^{t_1} d\tau W^{(i,0)}(t_1,\tau)\, W^{(j,0)}(t_2,\tau) + \\ + \sum_{k=0}^{i-1} \frac{\partial^k}{\partial t_1^k}[W^{(i-k-1,0)}(t_1,t_1)\, W^{(j,0)}(t_2,t_1)] + \\ + \sum_{k=0}^{n-1}\sum_{\ell=0}^{n-1} \Gamma_{k\ell}\, q_k^{(i)}(t_1)\, q_i^{(j)}(t_2) \;;\; t_2 > t_1 \end{cases} \tag{3-8}$$

and hence

$$\begin{aligned} J_{i,j}(t_2) &= \left\{ \sum_{\ell=0}^{j-1} \frac{\partial^\ell}{\partial t_2^\ell}[W^{(i,0)}(t_1,t_2)\, W^{(j-\ell-1,0)}(t_2,t_2)] \right\} t_1 \downarrow t_2 \\ &- \left\{ \sum_{k=0}^{i-1} \frac{\partial^k}{\partial t_1^k}[W^{(i-k-1,0)}(t_1,t_1)\, W^{(j,0)}(t_2,t_1)] \right\} t_1 \uparrow t_2 \end{aligned} \tag{3-9}$$

It should be noted that no initial condition terms appear in (3-7) or (3-9); i.e., the discontinuities of the $\Gamma_{YY}^{(i,j)}(t_1,t_2)$ depend only on the $W^{(k,j)}(t,\tau)$. If the $W^{(k,j)}(t,\tau)$ appearing in (3-9) are expressed in terms of the $a_i^{(k)}(t)$ and the $F_j^{(\ell)}(t)$ as developed in Section 2.3, then (3-9) becomes a second degree equation in the $F_j(t)$ and their derivatives. This form for the $J_{i,j}(t_2)$ will be examined in more detail in a later section.

3.4 A FINAL IDENTITY

Equations (3-10) can be considered as a set of simultaneous nonlinear integral equations in the $\beta_i(\tau)$ if the $p_i(t)$, $q_j(t)$, and Γ_{ij} are assumed

$$p_i(t) = \sum_{j=0}^{n-1} q_j(t) \int_0^t d\tau \beta_i(\tau)\beta_j(\tau) + \sum_{j=0}^{n-1} \Gamma_{ij}\, q_j(t) \quad ; \qquad i = 0, \ldots, n-1 \tag{3-10}$$

known. These integral equations can be converted into a set of simultaneous second degree differential equations in the $\beta_i(t)$ which are independent of the Γ_{ij}. To see this, the operator $\sum_{k=0}^{n-1} a_k(t)\, d^k/dt^k$ is applied to both sides of (3-10), [assuming, of course, the requisite differential properties for the $p_i(t)$ and $q_j(t)$] with the result

$$\sum_{k=0}^{n} a_k(t)\, p_i^{(k)}(t) = \sum_{j=0}^{n-1} \sum_{k=0}^{n} a_k(t) \frac{d^k}{dt^k} [q_j(t) \int_0^t d\tau \beta_i(\tau)\beta_j(\tau) + \Gamma_{ij}] \quad ; \qquad i = 0, \ldots, n-1 \tag{3-11}$$

Making use of the fact that the $q_j(t)$ satisfy (2-3) it is clear that the terms in (3-11) involving the Γ_{ij} vanish identically in t for all i,j. Expansion of the remaining terms in (3-11) according to the rule for differentiation of products gives

$$\sum_{k=0}^{n} a_k(t)\, p_i^{(k)}(t) = \sum_{j=0}^{n-1} \sum_{k=0}^{n} \sum_{\ell=0}^{k} \binom{k}{\ell} a_k(t)\, q_j^{(\ell)}(t) \times$$
$$\times \; [\int_0^t d\tau\, \beta_i(\tau)\, \beta_j(\tau)]^{(k-\ell)}; \qquad i = 0, \ldots, n-1 \tag{3-12}$$

Now (3-12) can be rewritten in the form

$$\sum_{k=0}^{n} a_k(t)\, p_i^{(k)}(t) \;=\; \sum_{j=0}^{n-1} \sum_{k=0}^{n} \sum_{\ell=0}^{k-1} \binom{k}{\ell} a_k(t)\, q_j^{(\ell)}(t) \;\times$$

$$\times\; [\beta_i(t)\beta_j(t)]^{(k-\ell-1)} + \sum_{j=0}^{n-1} \sum_{k=0}^{n} a_k(t)\, q_j^{(k)}(t) \;\times$$

$$\times\; [\int_0^t d\tau \beta_i(\tau)\beta_j(\tau)] \;;\; i = 0, \ldots, n-1 \qquad (3\text{-}13)$$

Again using the fact that the $q_j(t)$ are solutions of (2-3), it is clear that the second term on the right-hand side of (3-13) vanishes identically in t for every value of i and j. Further expansion of the terms $[\beta_i(t)\beta_j(t)]^{(k-\ell-1)}$ in (3-13) yields

$$\sum_{k=0}^{n} a_k(t)\, p_i^{(k)}(t) \;=\; \sum_{j=0}^{n-1} \sum_{k=0}^{n} \sum_{\ell=0}^{k-1} \sum_{m=0}^{k-\ell-1} \binom{k}{\ell} \binom{k-\ell-1}{m} \;\times$$

$$\times\; a_k(t)\, q_j^{(\ell)}(t)\beta_i^{(m)}(t)\beta_j^{(k-\ell-m-1)}(t) \;;$$

$$i = 0, \ldots, n-\ell \;;\; t \quad 0 \qquad (3\text{-}14)$$

The above identities form the desired set of second degree differential equations in the $\beta_i(t)$. They are obviously equivalent to (3-10) providing a proper set of initial conditions for the $\beta_i(t)$ and their derivatives is given.

3.5 A DISCUSSION OF BATKOV'S ERROR

As noted in Chapter 1, the second method described by Batkov in his paper only works as an algebraic method for a much more restricted class of covariance functions than claimed by Batkov. Using the results of Sections 2.3 and 3.3, this is easily demonstrated as follows.

Letting $n \geq 3$ and examining $J_{0,1}(t_2)$, it is found from (3-9) that

$$J_{0,1}(t_2) \;=\; W^2(t_2, t_2) \qquad (3\text{-}15)$$

which, by (2-26) and (2-21) yields

$$J_{0,1}(t_2) \;=\; F_1^2(t_2) \;=\; b_{n-1}^2(t_2) \qquad (3\text{-}16)$$

Similarly, from (3-9)

$$J_{0,2}(t_2) = W(t_2,t_2)\, W^{(1,0)}(t_2,t_2) + W^{(0,1)}(t_2,t_2)\, W(t_2,t_2) + W(t_2,t_2)\, W(t_2,t_2)^{(1)} \quad (3\text{-}17)$$

which by (2-26), (2-29), and (2-21) becomes

$$\begin{aligned} J_{0,2}(t_2) &= F_1(t_2)\, F_2(t_2) + -F_2(t_2) + F_1^{(1)}(t_2)\, F_1(t_2) + \\ &+ F_1(t_2)\, F_1^{(1)}(t_2) = 2F_1(t_2)\, F_1^{(1)}(t_2) \\ &= 2b_{n-1}(t_2)\, b_{n-1}^{(1)}(t_2) \end{aligned} \quad (3\text{-}18)$$

Note that $J_{0,2}(t_2)$ does not depend on $b_{n-2}(t_2)$ which is contrary to what Batkov claims [Batkov's $\Gamma_k(t_2,t_2)$ is equal to $J_{0,k}(t_2)$]. Further, from (3-9)

$$\begin{aligned} J_{0,3}(t_2) &= W(t_2,t_2)\, W^{(2,0)}(t_2,t_2) + W^{(0,1)}(t_2,t_2)\, W^{(1,0)}(t_2,t_2) \\ &+ W(t_2,t_2)\, W^{(1,0)}(t_2,t_2)^{(1)} + W^{(0,2)}(t_2,t_2)\, W(t_2,t_2) \\ &+ 2W^{(0,1)}(t_2,t_2)\, W(t_2,t_2)^{(1)} + W(t_2,t_2)\, W(t_2,t_2)^{(2)} \end{aligned} \quad (3\text{-}19)$$

which by (2-26) and (2-29)

$$\begin{aligned} J_{0,3}(t_2) &= F_1(t_2)\, F_3(t_2) + -F_2(t_2) + F_1^{(1)}(t_2)\, F_2(t_2) \\ &+ F_1(t_2)\, F_2^{(1)}(t_2) + -2F_2^{(1)}(t_2) + F_1^{(2)}(t_2) \\ &+ F_3(t_2)\, F_1(t_2) + 2\, -F_2(t_2) + F_1^{(1)}(t_2)\, F_1^{(1)}(t_2) \\ &+ F_1(t_2)\, F_1^{(2)}(t_2) \\ &= 2F_1(t_2)\, F_3(t_2) - F_2^2(t_2) - F_1^{(1)}(t_2)\, F_2(t_2) \\ &- F_2^{(1)}(t_2)\, F_1(t_2) + F_1^{(1)}(t_2)^2 + F_1(t_2)\, F_1^{(2)}(t_2) \end{aligned} \quad (3\text{-}20)$$

Making use of (2-21), it is a function of $b_{n-3}(t_2)$, $b_{n-2}(t_2)$ and its first derivative, and $b_{n-1}(t_2)$ and its first and second derivatives. Proceeding, it can be shown that a similar situation obtains for the higher order jumps, $J_{i,j}(t_2)$. This clearly demonstrates Batkov's error and shows that from the jumps one can obtain at best a set of simultaneous, nonhomogeneous, nonlinear differential equations in the $F_i(t_2)$ or $b_i(t_2)$. An independent set can be obtained from the jumps $J_{i,i+1}(t_2)$ where $i = 0, \ldots, n-1$ for example. However, because of their complexity and, hence, lack of utility, they are neither derived nor considered herein. Of course, when $b_i(t_2) \equiv 0$ for $0 < i \leq n - 1$, then $J_{i,j}(t_2) = 0$ for $i,j \leq n - 1$ and $J_{n-1,n}(t_2) = W^{(n-1,0)}(t_2,t_2)^2 = b_0^2(t_2)$ and Batkov's algebraic method works. However, the first method described by Batkov also works for this special case. Note that it has been assumed in the above that $\Gamma(t_1,t_2)$ is the covariance function of a process obtained from a *white noise* process by the physically realizable characterized by an n^{th} order differential equation of the form given in (2-1).

CHAPTER 4

EXACT SOLUTION OF THE SHAPING FILTER PROBLEM FOR SOME SPECIAL CASES

4.1 INTRODUCTION

When the function $\Gamma(t_1, t_2)$ introduced in Section 1.1 has certain special properties, the shaping filter problem can be resolved by fairly elementary methods. These special cases are studied in this chapter and the methods of solution are presented. It is assumed that $T = (0, T)$ and that $\Gamma(t_1, t_2)$ is separable and is given in its separable form.

4.2 THE "SIMPLEST" CASE

Probably the simplest case is where $\Gamma(t_1, t_2)$ is of the form

$$\Gamma(t_1, t_2) = \sum_{i=1}^{n} \sum_{j=1}^{n} \Gamma_{ij}\, q_i(t_1)\, q_j(t_2) \tag{4-1}$$

where it can be assumed without loss of generality that the $q_i(t)$ are linearly independent on T. In this case a set of necessary and sufficient conditions that $\Gamma(t_1, t_2)$ be a covariance function is that the matrix $[\Gamma_{ij}]$ be symmetric and nonnegative-definite. To prove the sufficiency of the conditions it is observed that if $[\Gamma_{ij}]$ is symmetric, then so is $\Gamma(t_1, t_2)$. Further, considering the expression for arbitrary z_k

$$\sum_{k=1}^{n} \sum_{\ell=1}^{n} \Gamma(t'_k, t'_\ell)\, z_k\, z_\ell = \sum_{i=1}^{n} \sum_{j=1}^{n} \Gamma_{ij} \sum_{k=1}^{n} \sum_{\ell=1}^{n} q_i(t'_k)\, q_j(t'_\ell)\, z_k\, z_\ell$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} \Gamma_{ij}\, y_i\, y_j \tag{4-2}$$

where $y_i = \sum_{k=1}^{n} q_i(t'_k)\, z_k$, it follows that $\Gamma(t_1, t_2)$ is nonnegative-definite if $[\Gamma_{ij}]$ is. This establishes the sufficiency. On the other hand, symmetry of $\Gamma(t_1, t_2)$ clearly implies symmetry of $[\Gamma_{ij}]$. Further, since the $q_i(t)$ are linearly independent on T, there exist at least n-values of $t \in T$, say t'_k, such that the matrix $[q_i(t'_k)]$ is nonsingular, for if not, the $q_i(t)$ would be linearly dependent on T. For arbitrary y_i, let z_k be defined by

$$[z] = [q_i(t'_k)]^{-1}[y] \tag{4-3}$$

where $[z]^T = [z_1 \ . \ . \ . \ z_n]$ and $[y]^T = [y_1 \ . \ . \ . \ y_n]$. Hence, the expression

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \Gamma_{ij}\, y_i y_j \;=\; \sum_{k=1}^{n} \sum_{\ell=1}^{n} \Gamma(t'_k\,,\, t'_\ell)\, z_k\, z_\ell \geq 0 \tag{4-4}$$

holds for arbitrary y_i, the inequality following the nonnegative-definiteness of $\Gamma(t_1, t_2)$. This establishes the necessity of the conditions.

Now given that $[\Gamma_{ij}]$ is symmetric and nonnegative-definite, the process whose sample functions are of the form

$$X(t) \;=\; \sum_{i=1}^{n} A_i\, q_i(t) \tag{4-5}$$

where the A_i are random variables and $E\, A_i A_j = \Gamma_{ij}$ has $\Gamma(t_1, t_2)$ for its covariance function. If the $q_i(t)$ have continuous n^{th} order derivatives on T and if their Wronskian does not vanish on T, then the process $\{X(t)\}$ can be generated by a physically realizable system characterized by (2-3) with random initial conditions and no input; i.e., as the *transient* response. The $a_i(t)$ in (2-3) can be found by (2-31) and the covariances of the initial conditions are given by

$$[\Gamma'_{ij}] \;=\; [W]^{T}[\Gamma_{ij}]\,[W] \tag{4-6}$$

where $\Gamma'_{ij} \;=\; E[X^{(i)}(0)\; X^{(j)}(0)]$ and $[W]$ is the matrix whose i,j element is $q_i^{(j)}(0)$. Of course $\{X(t)\}$ can always be generated by the physically realizable system composed of n-function generators, n-multipliers with random magnitudes A_i, and a summing amplifier. This resolves the shaping filter problem for this case.

While it is not directly related to the shaping filter problem, it is interesting and worthwhile to consider the predictability of processes whose covariance functions are of the form given in (4-1). In view of (4-5), it is not surprising to find that such a process is essentially predictable exactly. In fact, if the sample functions of the process are indeed those given in (4-5), then being able to predict the *future* values of a sample function $X(t)$ is just a matter of being able to find the values of the A_i for the particular sample function $X(t)$ of interest in terms of the observed values of $X(t)$. Now if the $q_i(t)$ are linearly independent on some interval observation $I \subset T$, then, as before, there exist at least n-values of $t \in T$, say t'_i, such that the matrix $[q_j(t'_i)]$ is nonsingular. Hence, given the observed values of $X(t)$ at the points t'_i, the A_i can be determined exactly by the expression

$$\begin{bmatrix} A_1 \\ \vdots \\ A_n \end{bmatrix} = \begin{bmatrix} q_1(t'_1)\ldots & q_n(t'_1) \\ \vdots & \vdots \\ q_n(t'_n)\;\ldots & q_n(t'_n) \end{bmatrix}^{-1} \begin{bmatrix} X(t'_1) \\ \vdots \\ X(t'_n) \end{bmatrix} \tag{4-7}$$

and thus the predicted value $\hat{X}(t)$ given by

$$\hat{X}(t) = [q_1(t) \cdots q_n(t)] \begin{bmatrix} q_1(t_1') \cdots & q_n(t_1') \\ \vdots & \vdots \\ q_n(t_n') \cdots & q_n(t_n') \end{bmatrix}^{-1} \begin{bmatrix} X(t_1') \\ \vdots \\ X(t_n') \end{bmatrix} \tag{4-8}$$

equals $\hat{X}(t)$ for all $t \in T$. Since this will be true for *every* sample function of the process, it follows that $X(t) = \hat{X}(t)$ with probability 1 for all $t \in T$. Even if the sample functions of the process are not known to be of the form given in (4-5), as long as the covariance function of the process is the same, the predicted value $\hat{X}(t)$ still equals $X(t)$ with probability 1 for all $t \in T$. To prove this, it is sufficient to show that $E|X(t) - \hat{X}(t)|^2 = 0$. Now

$$E|X(t) - \hat{X}(t)|^2 = \Gamma(t,t) - 2E\,\hat{X}(t)X(t) + E\,\hat{X}^2(t) \tag{4-9}$$

Computing $E\,\hat{X}(t)\,X(t)$ it is found that

$$E\,\hat{X}(t)\,X(t) = [q_1(t) \cdots q_n(t)] \begin{bmatrix} \\ q_j(t_k') \\ \\ \end{bmatrix}^{-1} \begin{bmatrix} \Gamma(t_1', t) \\ \vdots \\ \Gamma(t_n', t) \end{bmatrix} \tag{4-10}$$

and from (4-1) it follows that

$$\begin{bmatrix} \Gamma(t_1', t) \\ \vdots \\ \Gamma(t_n', t) \end{bmatrix} = [q_j(t_i')]\,[\Gamma_{jk}] \begin{bmatrix} q_1(t) \\ \vdots \\ q_n(t) \end{bmatrix} \tag{4-11}$$

Substituting (4-11) into (4-10) and making use of (4-1), it follows that $E\,\hat{X}(t)\,X(t) = \Gamma(t, t)$. Similarly, it is also easily shown that $E\,\hat{X}^2(t) = \Gamma(t, t)$ and hence $E|X(t) - X(t)|^2 = 0$ for all $t \in T$. Thus it has been shown that if the $q_i(t)$ are linearly independent on an observation interval I, then the process is predictable with probability 1 for all $t \in T$ by a linear combination of n-observations of $X(t)$ on the interval I. Since the $q_i(t)$ are linearly independent on T, this clearly implies that, if a process has a covariance function of the form given in (4-1), then its sample functions have the representation (4-5) where equality holds with probability 1.

On the other hand, suppose that the $q_i(t)$ are linearly dependent on an observation interval I. Then the sample functions of the process have the representation

$$X(t) = \sum_{i=1}^{r} B_i \, q_i(t) \quad , \quad r < n \tag{4-12}$$

for $t \in I$ where the $q_i(t)(i = 1,\ldots,r)$ are linearly independent on I and the B_i are linear combinations of the A_i; the equality in (4-12) holding with probability 1. Now, suppose the process is predictable with probability 1 for all $t \in T$ in terms of a linear combination of values of $X(t)$ for $t \in I$. Then the representation (4-12) holds for all $t \in T$. But, providing $[\Gamma_{ij}]$ is positive-definite, this contradicts the assumption of linear independence of the $q_i(t)(i=1,\ldots,n)$ on T. This result is the converse of that of the preceding paragraph for the case where $[\Gamma_{ij}]$ is positive-definite. Since, if $[\Gamma_{ij}]$ is not positive-definite, the number of terms in its covariance function can be reduced to the point where it is by defining new $q_i(t)$ as linear combinations of the original $q_i(t)$, the assumption of nonnegative-definiteness of $[\Gamma_{ij}]$ entails no loss of generality.

An interesting generalization of the above is obtained by letting $n=\infty$. Then $\Gamma(t_1, t_2)$ is of the form

$$\Gamma(t_1, t_2) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \Gamma_{ij} \, q_i \, (t_1) \, q_j \, (t_2) \tag{4-13}$$

Of course, in order for (4-13) to be meaningful, the mode of convergence of the double series must be specified. A natural mode of convergence is pointwise on $T \times T$ and this shall be the mode specified. Again, the $q_i(t)$ are, without loss of generality, assumed to be independent on T. Using the argument used before and letting $n=\infty$, it is easily shown that if the matrix $[\Gamma_{ij}]$ $i, j = 1, \ldots, m$ is symmetric and nonnegative-definite for all m, then $\Gamma(t_1, t_2)$ is a covariance function. Now consider the processes $\{X_n(t)\}$ whose sample functions are of the form

$$X_n(t) = \sum_{i=1}^{n} A_i \, q_i \, (t) \tag{4-14}$$

where the A_i are random variables satisfying $E \, A_i \, A_j = \Gamma_{ij}$. Now, since for all t

$$E|X_n(t) - X_m(t)|^2 = \sum_{i=m+1}^{n} \sum_{j=m+1}^{n} \Gamma_{ij} \, q_i(t) \, q_j(t) \xrightarrow[n,m \to \infty]{} 0 \tag{4-15}$$

the convergence to zero in (4-15) being a consequence of the convergence of (4-13), the $X_n(t)$ converge in the mean to some limit sample functions $X(t)$ of a limit process $\{X(t)\}$. Further, it follows that the covariance function of $\{X(t)\}$ exists and is that given in (4-13). Formally this can be stated as

$$X(t) = \underset{n \to \infty}{\text{l.i.m.}} \, X_n(t) = \sum_{i=1}^{\infty} A_i \, q_i(t) \tag{4-16}$$

This resolves the shaping filter problem for this generalized case.

Turning now to consideration of the predictability of processes whose covariance functions are of the form given in (4-13), a result analogous to that given above for $n < \infty$ is fairly easily established providing certain additional mild requirements are satisfied on the observation interval I. To see how to proceed, suppose for the moment that the sample functions of the process are of the form given in (4-16). Then formally, if there existed a set of functions $f_i(t)$ such that $\int_I f_i(t)\, q_j(t)\, dt = \delta_{ij}$, the A_i could be found by multiplying both sides of (4-16) by the $f_i(t)$ respectively and integrating over the observation interval I. Once the A_i have been determined, the predictability follows immediately. Now the additional requirements mentioned above are just those needed to rigorously justify this procedure in general [that is, even when it is not known a priori that the sample functions of the process are of the form given in (4-16)]. This motivation leads formally to the consideration of

$$\hat{X}_n(t) = \sum_{i=1}^{n} q_i(t) \int_I f_i(t)\, X(t)\, dt \tag{4-17}$$

as a predicted value of $X(t)$ where hopefully the error of prediction goes to zero as $n \to \infty$.

Specifically, assume that the $q_i(t)$ are linearly independent on I in the sense that for all finite n there exists no set of constants a_i not all zero such that $\sum_{i=1}^{n} a_i\, q_i(t) = 0$ almost everywhere on I* and that $\int_I q_i^2(t)dt < \infty$ for all i. Then as shown in Appendix 1 there exists a set of functions $f_i(t)$ defined on I such that $\int_I f_i^2(t)\, dt < \infty$ for all i and $\int_I f_i(t)\, q_j(t)\, dt = \delta_{ij}$. Finally, assume that one of the following statements is true: Either (i) $\left|\sum_{i=1}^{n} \sum_{j=1}^{\infty} \Gamma_{ij}\, q_i(t_1)\, q_j(t_2)\right| \leq S_1(t_1)$ almost everywhere on I for all n and all $t_2 \epsilon I$ and $\left|\sum_{i=1}^{n} \Gamma_{ij}\, q_i(t_1)\right| \leq S_2(t_1)$ almost everywhere on I for all n, j where $S_1(t_1)$ and $S_2(t_2)$ are integrable on I, or (ii) $\int_I \Gamma^2(t_1, t_2)\, dt_1 < \infty$ and the series $\sum_{i=1}^{n} \sum_{j=1}^{\infty} \Gamma_{ij}\, q_i(t_1)\, q_j(t_2)$ converges weakly to $\Gamma(t_1, t_2)$ on I for all $t_2 \epsilon I$ and $\int_{I \times I} \Gamma^2(t_2, t_1)\, dt_1\, dt_2 < \infty$ and the series $\sum_{i=1}^{n} \sum_{j=1}^{n} \Gamma_{ij}\, q_i(t_1)\, q_j(t_2)$ converges weakly to $\Gamma(t_1, t_2)$ on $I \times I$.** Under these assumptions consider the expression

$$E|X(t) - \hat{X}_n(t)|^2 = \Gamma(t,t) - 2E\hat{X}_n(t)\, X(t) + E\, \hat{X}_n^2(t) \tag{4-18}$$

---

*Note that this definition of linear independence differs slightly from the classical definition used above for the case where $n < \infty$.

**Note that convergence in the mean implies weak convergence, [18, p.175].

Computing $E\,\hat{X}_n(t)\,X(t)$ there results

$$E\,\hat{X}_n(t)\,X(t) = \sum_{i=1}^{n} q_i(t) \int_I f_i(\tau)\Gamma(\tau,t)\,d\tau$$

$$= \sum_{i=1}^{n} q_i(t) \sum_{\ell=1}^{\infty} \sum_{j=1}^{\infty} \Gamma_{j\ell}\, q_j(t) \int_I f_i(\tau)\, q_\ell(\tau)\,d\tau$$

$$= \sum_{i=1}^{n} \sum_{y=1}^{\infty} \Gamma_{ij}\, q_i(t)\, q_j(t) \xrightarrow[n\to\infty]{} \Gamma(t,t) \qquad (4\text{-}19)$$

the interchange of order of summation and integration being justified by the assumption of bounded convergence or weak convergence of the relevant series. Similarly, computing $E\,\hat{X}_n^2(t)$ there results

$$E\,\hat{X}_n^2(t) = \sum_{i=1}^{n} \sum_{i=1}^{n} q_i(t)\, q_j(t) \int_I d\tau_1 \int_I d\tau_2\, f_i(\tau_1)\, f_j(\tau_2)\Gamma(\tau_1,\tau_2)$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} q_i(t)\, q_j(t) \int_I d\tau_1 \sum_{k=1}^{\infty} \sum_{\ell=1}^{\infty} \Gamma_{k\ell}\, f_i(\tau_1)\, q_k(\tau_1)$$

$$\int_I d\tau_2\, f_j(\tau_2)\, q_\ell(\tau_2)$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} \Gamma_{ij}\, q_i(t)\, q_j(t) \xrightarrow[n\to\infty]{} \Gamma(t,t) \qquad (4\text{-}20)$$

Hence,

$$E|X(t) - \hat{X}_n(t)|^2 \xrightarrow[n\to\infty]{} 0, \text{ all } t \in T \qquad (4\text{-}21)$$

which proves that under the above mild assumptions, processes whose covariance functions are of the form given in (4-13) are predictable in the mean. As before, linear independence of the $q_i(t)$ is of crucial importance; it being used in the proof of the existence of $f_i(t)$ which satisfy $\int_I f_i(t)\, q_j(t)\,dt = \delta_{ij}$.

The above result includes as a special case, two classes of processes which are known to be predictable in the mean for any nondegenerate observation interval $I \subset T$; namely, those with analytic covariance functions and those stationary processes whose spectral distribution functions are step functions. Finally, it is interesting to note that if one has expanded what Wiener calls an innovation process; e.g., a stationary process with absolutely continuous spectral distribution function, in an infinite series over an interval T by the Karhunen-Loeve expansion theorem or by any other method leading to reasonably convergent series for $\Gamma(t_1,t_2)$, then while the $q_i(t)$ in the expansion are linearly independent on T, they cannot be linearly independent on any proper subinterval of T. If they were, the process would be predictable in the mean in terms of values on the subinterval which contradicts the assumption that the process is an innovation process.

4.3 THE "ALMOST STATIONARY" CASE

Another case for which the shaping filter problem can be resolved by elementary methods is that where $\Gamma(t_1,t_2)$ is of the form

$$\Gamma(t_1,t_2) = \sum_{i=1}^{n} D_i e^{-a_i|t_1-t_2|} + \sum_{i=1}^{n}\sum_{j=1}^{n} C_{ij}\, e^{-a_i t_1}\, e^{-a_j t_2} \tag{4-22}$$

where $R_e a_i > 0$ and the a_i and D_i are real or occur in complex conjugate pairs. Here, assuming for the moment that $\Gamma(t_1,t_2)$ is a covariance function, the non-stationary character of $\Gamma(t_1,t_2)$ arises solely due to *transients* which die out as $t_1,t_2 \to \infty$, the corresponding process being asymptotically stationary.

Now a sufficient condition that $\Gamma(t_1,t_2)$ be a covariance function is that the Fourier Transform of $\sum_{i=1}^{n} D_i\, e^{-a_i|\tau|}$ be everywhere positive and the matrix $[C_{ij}]$ be nonnegative-definite. In this case one can find, by the usual method of factoring the Fourier Transform of $\sum_{i=1}^{n} D_i\, e^{-a_i|\tau|}$, a weighting function for a shaping filter of the form

$$W(\tau) = \begin{cases} \sum_{i=1}^{n} d_i\, e^{-a_i\tau} & , \tau \geq 0 \\ 0 & , \tau < 0 \end{cases} \tag{4-23}$$

whose corresponding inverse is stable. If a *white noise* process is applied to the shaping filter at $t = 0$ with the shaping filter assumed to be at rest at $t = 0$, then the covariance function $\Gamma_1(t_1,t_2)$ of the output process is

$$\Gamma_1(t_1,t_2) = \sum_{i=1}^{n} D_i\, e^{-a_i|t_1-t_2|} - \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{d_i\, d_j}{a_i + a_j} e^{-a_i\, t_1 - a_j\, t_2} \tag{4-24}$$

Note that the function $\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{d_i\, d_j}{a_i + a_j}\, e^{-a_i\, t_1}\, e^{-a_j\, t_2}$ is a covariance function on $T \times T$. This follows from the fact that if a *white noise* process is applied to the shaping filter over the remote past and then removed at $t = 0$, then the covariance function of the resulting *transient* output process for $t \geq 0$ will be $\sum_{i=1}^{n}\sum_{j=1}^{n} \frac{d_i\, d_j}{a_i + a_j}\, e^{-a_i\, t_1}\, e^{-a_j\, t_2}$. Subtracting (4-24) from (4-22) there results

$$\Gamma_2(t_1,t_2) = \sum_{i=1}^{n}\sum_{j=1}^{n} \Gamma_{ij}\, e^{-a_i\, t_1}\, e^{-a_j\, t_2} \tag{4-25}$$

which, by the assumptions made on $[C_{ij}]$, is a covariance function. Making use of the results of Section 4.2, this establishes that with a *white noise* source and a set of random initial conditions which are uncorrelated with the *white noise* source, a process whose covariance function is of the form given in (4-22), can be generated by the system whose weighting function is that given in (4-23). Of course, even if the matrix $[C_{ij}]$ is not nonnegative-definite, the result still holds as long as $[\Gamma_{ij}]$ is nonnegative-definite.

4.4 THE "NONDEGENERATE" CASE

The final case where the shaping filter problem can be resolved by elementary methods is based on consideration of (3-3). Suppose that one were given a function $\Gamma(t_1, t_2)$ expressed in the form given in (3-4) where each of the $p_i(t)$ is a sum of at least n-terms. Suppose further that upon division of $p_i(t)$ by $q_j(t)$ and $p_j(t)$ by $q_i(t)$, it happens that the resulting sums have a term in common. Then by (3-3) it might be assumed that this term is $\int_0^t dt \beta_i(t)\beta_j(t) + \Gamma_{ij}$. If this happens for all $j \neq i$, then the one term of $p_i(t)$ which, after the appropriate division, did not appear to be a term which $p_i(t)$ had in common with some $p_j(t)$ might be assumed to be $q_i(t)[\int_0^t \beta_i^2(\tau)\, d\tau + \Gamma_{ii}]$. If it were, then $\beta_i(t)$ could be found by dividing it by $q_i(t)$, differentiating with respect to t, and taking the square root of the resulting derivative. Once the $\beta_i(t)$ have been determined, the determination of the Γ_{ij} is obviously trivial.

Whenever the procedure sketched above works, it will be said that one is dealing with a nondegenerate case. At first glance it might appear that this is indeed a special case since, for example, stationary processes and almost stationary processes are clearly degenerate cases. However, there appears to be a reasonably large class of <u>nonstationary</u> processes which lead to nondegenerate cases. For example, consider the case where $\Gamma(t_1, t_2)$ is

$$\Gamma(t_1, t_2) = \begin{cases} t_1[t_2^{6/5} + t_2^{9/7}] + t_1^2[t_2^{8/7} + t_2^{11/9}] & ; \ t_1 \geq t_2 \geq 0 \\[2ex] t_2[t_1^{6/5} + t_1^{9/7}] + t_2^2[t_1^{8/7} + t_1^{11/9}] & ; \ t_2 \geq t_1 \geq 0 \end{cases} \tag{4-26}$$

Here

$$\begin{aligned} q_1(t) &= t \ , \quad p_1(t) = t^{6/5} + t^{9/7} \\ q_2(t) &= t^2 \ , \quad p_2(t) = t^{8/7} + t_1^{11/9} \end{aligned} \tag{4-27}$$

Upon dividing $p_1(t)$ by t^2 and $p_2(t)$ by t, it is observed that $t^{7/7}$ is a common term. Hence, $t^{6/5}$ appears to be $q_1(t)[\int_0^t d\tau \beta_1^2(\tau) + \Gamma_{11}]$. Proceeding under this assumption, it is found that $\beta_1(t) = \pm\ t^2$ and $\Gamma_{11} = 0$. Similarly, it is found that $\beta_2(t) = \pm\ t^4$ and $\Gamma_{22} = 0$. Letting $t^{9/7} = q_2(t)[\int_0^t d\tau \beta_1(\tau)\beta_2(\tau)] + \Gamma_{12}$, it is found that this relationship is indeed satisfied if $\Gamma_{12} = 0$ and similarly for $t^{8/7}$. Hence, the weighting function of an appropriate shaping filter is

$$
\begin{aligned}
W(t,\tau) &= t\tau^2 + t^2\tau^4 \ , \quad t \geq \tau \geq 0 \\
&= \quad 0 \qquad\quad , \quad t < \tau
\end{aligned}
\qquad (4\text{-}28)
$$

There does not appear to be any simple criterion for determining whether a given function $\Gamma(t_1, t_2)$ is the covariance function for a nondegenerate case short of attempting to carry out the above procedure. Since the procedure is rather simple and direct, it is reasonable to just proceed as if it were a nondegenerate case and, if it fails at some step, then one concludes that he is dealing with a degenerate case and a more complex procedure is required. If it works, the shaping filter problem has been resolved rather easily. As the example shows, sometimes nonstationary cases are easier than stationary cases.

Note that when it works, the above procedure always yields a *physically realizable* shaping filter. Also note that after the Γ_{ij} have been determined, it is necessary to check and make sure that the matrix $[\Gamma_{ij}]$ is nonnegative-definite and, hence, represents the covariance matrix of a set of random initial conditions on the shaping filter.

CHAPTER 5

SOME FUNDAMENTAL RESULTS ON EXISTENCE AND UNIQUENESS

5.1 INTRODUCTION

As is well known, *physically realizable* shaping filters do not, in general, exist for processes with arbitrary covariance functions. For example, as pointed out in Section 1.2.1 for the stationary case, physical realizability of the shaping filter requires that the spectral distribution function be absolutely continuous and satisfy the Paley-Wiener criterion given in (1-13). Apparently no simple criterion analogous to that of Paley and Wiener has been developed for the general nonstationary case. This is not too surprising considering the difficulty of the problem. In this chapter, the question of the existence of *physically realizable* shaping filters is investigated for the class of separable covariance functions, and for this class, it is shown that, providing one remarkably simple requirement is met, a physically *realizable shaping* filter does indeed exist. In addition, the question of uniqueness of the shaping filter is also discussed. The restriction to the class of separable covariance functions certainly seems reasonable in view of the fact that in this case the resulting shaping filter is usually rather easily realized physically. This is, of course, of extreme importance in engineering applications.

It is rather interesting and enlightening to examine the treatment of this question for the nonstationary case by the authors whose work is summarized in Chapter 1. Darlington was apparently well aware of the problem and did provide answers for two rather restrictive cases. They were restrictive in the sense that he *assumed physical realizability* of the underlying signal and noise-shaping filters and either periodicity or regularity at ∞ of the corresponding differential equation. His answers clearly leave much to be desired. Batkov simply avoided the problem by making implicit the assumption that the covariance function was of the required form (a trick quite commonly used in writing technical papers). One wonders if he was even aware of the problem. Leonov wasn't concerned about physical realizability of the shaping filter (and didn't discuss or obtain it) because it wasn't required for his application. Actually, for his application, he didn't need shaping filters at all.

Finally, it should be noted that physical realizability of the shaping filter is automatically achieved for the special cases discussed in Chapter 4.

5.2 CASTING AND RECASTING THE PROBLEM

Neglecting initial conditions for the moment, establishing the existence of a physically realizable shaping filter amounts to establishing the existence of a solution of the nonlinear Volterra integral equation of the first kind

$$\Gamma(t_1,t_2) = \int_0^{t_2} d\tau \; W(t_1,\tau)\; W(t_2,\tau)\; d\tau \;, \quad t_1 \geq t_2 \geq 0 \tag{5-1}$$

When the covariance function is separable; i.e.,

$$\Gamma(t_1,t_2) = \sum_{i=1}^{n} q_i(t_1)\; p_i(t_2) \;;\; t_1 \geq t_2 \geq 0 \tag{5-2}$$

it is reasonable in view of Chapters 2 and 3 to consider solutions of the form

$$W(t,\tau) = \begin{cases} \sum_{i=1}^{n} q_i(t)\beta_i(\tau) \;;\; t \geq \tau \geq 0 \\ 0 \qquad ;\; t < \tau \end{cases} \tag{5-3}$$

In this case, the integral equation becomes

$$\sum_{i=1}^{n} q_i(t_1)\; p_i(t_2) = \sum_{i=1}^{n} q_i(t_1) \sum_{j=1}^{n} q_j(t_2) \int_0^{t_2} d\tau \; \beta_i(\tau)\; \beta_j(\tau) \tag{5-4}$$

and upon making use of the linear independence of the $q_i(t_1)$ and adding on initial conditions terms, there results

$$p_i(t) = \sum_{j=1}^{n} q_j(t) \left[\int_0^{t} d\tau \beta_i(\tau)\beta_j(\tau) + \Gamma_{ij} \right] \;;\; t \geq 0, \quad i = 1;\ldots, n \tag{5-5}$$

Thus for the case of separable covariance functions, the problem has been reduced to establishing the existence of a solution of the simultaneous nonlinear Volterra integral equations of the first kind given in (5-5). This certainly represents a reduction over (5-1) since (5-1) actually represents an infinite set of simultaneous integral equations, one for each value of t_1.

The usual procedure in the study of Volterra integral equations of the first kind is to first convert the integral equation into an integral equation of the second kind and then apply the standard techniques known for Volterra integral equations of the second kind. For linear equations, this conversion is easily carried out, [25,p.16]. That such a procedure can also be carried out for the equations in (5-5) is perhaps not obvious, but nevertheless it can be accomplished as follows.

Upon differentiating the equations in (5-5), there results

$$p_i^{(1)}(t) = \sum_{j=1}^{n} q_j^{(1)}(t) \left[\int_0^t d\tau \beta_i(\tau)\beta_j(\tau) + \Gamma_{ij} \right] + \beta_i(t) \sum_{j=1}^{n} q_j(t)\beta_j(t) \tag{5-6}$$

Now examination of (5-6) shows that the multiplier of $\beta_i(t)$ is the same for all i; namely, $\sum_{j=1}^{n} q_j(t)\beta_j(t)$. Let $k(t) = \sum_{j=1}^{n} q_j(t)\beta_j(t)$. The only problem is that $k(t)$ is unknown. If $k(t)$ were known and nonzero for all $t \geq 0$, then the desired conversion to integral equations of the second kind would be complete upon division by $k(t)$. While at first glance it may appear that, since $k(t)$ involves the unknown $\beta_i(t)$, there is no hope of being able to determine it, it can be determined. Solving (5-6) for $\beta_i(t)$ it is found that

$$\beta_i(t) = k^{-1}(t) \left\{ p_i^{(1)}(t) - \sum_{j=1}^{n} q_j^{(1)}(t) \left[\int_0^t d\tau \beta_i(\tau)\beta_j(\tau) + \Gamma_{ij} \right] \right\} \tag{5-7}$$

Substituting back into (5-6) there results

$$\begin{aligned} p_i^{(1)}(t) - \sum_{j=1}^{n} q_j^{(1)}(t) \int_0^t d\tau \beta_i(\tau)\beta_j(\tau) + \Gamma_{ij} &= \\ = k^{-2}(t) \left\{ p_i^{(1)}(t) - \sum_{j=1}^{n} q_j^{(1)}(t) \left[\int_0^t d\tau \beta_i(\tau)\beta_j(\tau) + \Gamma_{ij} \right] \right\} &\times \\ \times \left\{ \sum_{j=1}^{n} q_j(t) \left(p_j^{(1)}(t) - \sum_{k=1}^{n} q_j(t)\; q_k^{(1)}(t) \left[\int_0^t d\tau \beta_j(\tau)\beta_k(\tau) + \Gamma_{jk} \right] \right) \right\} & \end{aligned} \tag{5-8}$$

which, upon cancelling and rearranging, yields

$$\begin{aligned} k^2(t) = \sum_{j=1}^{n} q_j(t)\; p_j^{(1)}(t) - \sum_{k=1}^{n} q_k^{(1)}(t) \sum_{j=1}^{n} q_j(t) &\times \\ \times \left[\int_0^t d\tau \beta_j(\tau)\beta_k(\tau) + \Gamma_{jk} \right] & \end{aligned} \tag{5-9}$$

But the second term in (5-9) is just $p_k(t)$ and hence

$$k(t) = \pm\sqrt{\sum_{j=1}^{n}\left[q_j(t)\, p_j^{(1)}(t) - q_j^{(1)}(t)\, p_j(t)\right]} \tag{5-10}$$

Thus $k(t)$ *can be determined* from $\Gamma(t_1,t_2)$. Making use of (3-4) and (3-6), $k(t)$ can be expressed in the form

$$k(t) = \pm\sqrt{\Gamma^{(0,1)}(t,\tau)\big|_{\tau\uparrow t} - \Gamma^{(0,1)}(t,\tau)\big|_{\tau\downarrow t}} = \pm\sqrt{J_{0,1}(t)} \tag{5-11}$$

It is quite important to note that if the $\beta_i(t)$ are to be real then (5-7) together with (5-11) requires that $J_{0,1}(t) \geq 0$ for all $t \geq 0$. This completes the conversion of the integral equations of the first kind given in (5-5) to integral equations of the second kind as given in (5-7).

Of course, there still remains the problem of what to do in case $k(t) \equiv 0$. In this case, the equations given in (5-6) reduce to a set of integral equations of the first kind of the form given in (5-5) with $q_i(t)$ and $p_i(t)$ replaced by $q_i^{(1)}(t)$ and $p_i^{(1)}(t)$ respectively. Hence, the logical thing to do is to re-apply the conversion procedure used in the previous paragraph on (5-5). When this is done one obtains

$$\beta_i(t) = k_1^{-1}(t)\left\{p_i^{(2)}(t) - \sum_{j=1}^{n} q_j^{(2)}(t)\left[\int_0^t d\tau\beta_i(\tau)\beta_j(\tau) + \Gamma_{ij}\right]\right\} \tag{5-12}$$

where

$$k_1(t) = \pm\sqrt{\sum_{i=1}^{n} q_i^{(1)}(t)\, p_i^{(2)}(t) - q_i^{(2)}(t)\, p_i^{(1)}(t)} = \pm\sqrt{J_{1,2}(t)} \tag{5-13}$$

For the same reason as before, it is required that $J_{1,2}(t) \geq 0$ for all $t \geq 0$. Naturally, if $k_1(t) \equiv 0$, then one re-applies the procedure to the new equations etc.

When $k(t) = 0$ [or $k_1(t) = 0$, etc.] for some values of t but not identically, then one is dealing with a more complicated type of integral equation which Picard called an equation of the third kind. Such cases have been studied for linear equations by Lalesco [26].

The problem now has been reduced to establishing the existence of a solution of the integral equations of the second kind given in (5-7). To do this, use will be made of some results due to T. Sato [27]. Since Sato makes use of Schauder's fixed point theorem in his treatment of existence questions, a short discussion of fixed point theorems is in order.

5.3 DISCUSSION OF FIXED POINT THEOREMS

The basic idea underlying fixed point theorems can be nicely demonstrated by the following simple example [28, p.118]. Let C be the set $\{X : 0 \leq X \leq 1\}$ and let $\sigma(X)$ be a continuous, single-valued transformation of C into itself (i.e., $\sigma(x)$ is a continuous, single-valued function defined on $[0,1]$ for which $\sigma(x) \epsilon [0,1]$ for all $x \epsilon [0,1]$). Then there exists an $x_0 \epsilon C$ such that $x_0 = \sigma(x_0)$. x_0 is called a fixed point for the transformation $\sigma(x)$. The truth of this result is obvious from Figure 2.

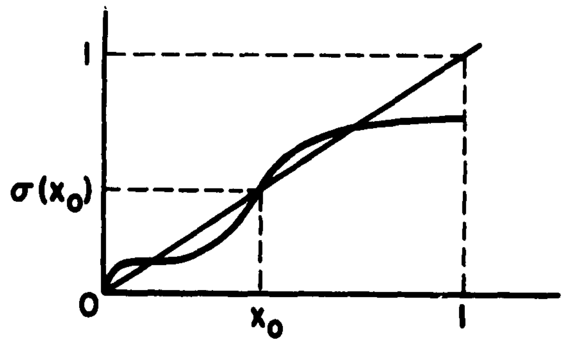

FIGURE 2

It is also obvious that x_0 may be either 1 or 0 and that it is not necessarily unique (there are four fixed points in Figure 2).

The generalization of this simple result to more general sets C in more general underlying topological spaces has led to the development of rather powerful (fixed point) theorems for establishing the existence of solutions (fixed points) of functional equations in general and integral equations in particular. For integral equations C becomes a class of functions and σ is an integral operator; e.g., the right-hand side of (5-7); and asserting the existence of a fixed point for σ is clearly equivalent to asserting the existence of a solution of the corresponding integral equation. One of the most general fixed point theorems and the one apparently used by Sato was proven by Schauder and can be stated as follows: [29, p.260].

> **Schauder's Theorem:** Let C be a nonempty, compact, convex set from a locally convex space X and let σ be a continuous, single-valued transformation of C into C. Then there exists an $x_0 \epsilon C$ such that $\sigma(x_0) = x_0$.

In applying Schauder's Theorem, the essential problem is, of course, to find an appropriate class C for the problem at hand.

It is interesting to note that the requirements of compactness and convexity of C stated in the theorem could have been anticipated on the basis of the simple example given above.

5.4 RESOLUTION OF THE EXISTENCE PROBLEM

Because Sato's treatment of the existence of solutions of the integral equations he studied is rather sketchy in nature, a development of it is given below. The development is essentially that given by Sato except that many of the *obvious* (to Sato!) steps are filled in, an error is corrected, and the appropriate space X and set C are clearly defined. The theorem resulting from this development is then applied to the set of integral equations given in (5-7) and the existence of *physically realizable* shaping filters is thereby deduced.

The following notation will be useful

I_r: the closed interval $0 \leq x \leq r$

Δ_r: the closed domain $0 \leq t \leq x \leq r$ in the plane (x,t)

D: the closed domain in the space $(x,t,u_1,\ldots,u_n)$ defined by $(x,t)\in \Delta_r$, $|u_i - f_i(x)| \leq \rho$ where the $f_i(x)$ are continuous functions on I_r and $\rho/2 > \max_i [\max_{I_r} f_i(x) - \min_{I_r} f_i(x)] \geq 0$

(For any $\rho > 0$ there is obviously an I_r, $r > 0$ such that the latter inequality is satisfied.)

Now consider the set of integral equations

$$u_i(x) = f_i(x) + \int_0^x K_i[x,t,u_1(t), \ldots, u_n(t)]dt \tag{5-14}$$

where $K_i[x,t,u_1, \ldots, u_n]$ is continuous on D. Hence, there exists an M such that $|K_i[x,t,u_1, \ldots, u_n]| \leq M$ for all i and all $(x,t,u_1, \ldots, u_n)\in D$. Let $u(x)$ be the vector-valued function whose components are $u_i(x)$. Having made the above assumptions, the problem now is to find a space X and a set C which satisfy the hypotheses of Schauder's Theorem.

Let X be the space of all continuous vector functions $u(x)$ on I_r and let X have the topology of uniform convergence on compact (i.e., on all compact subsets of I_r) [30, p.226]. Then X is clearly a locally convex space. (For a definition of a locally convex space see [29, p.257].) Now, let F be any set of vector functions $u(x) \epsilon X$ which satisfy the conditions

$$|u_i(x) - f_i(x)| \leq \rho/2 \; ; u_i(0) = f_i(0) \tag{5-15}$$

on $I_{r'}$ where $r' = \min(r, \rho/2M)$. Then the right-hand side of (5-14), considered as a transformation σ, obviously transforms F into a set $\overline{F}$ of vector functions $\overline{u}(x) \epsilon X$ which also satisfy (5-15) on $I_{r'}$. Furthermore, the set $\overline{F}$ is seen to be equicontinuous on $I_{r'}$ (in fact, $\overline{F}$ is uniformly equicontinuous on I_r). Hence, for every $\epsilon \geq 0$, there exists a $\delta(\epsilon) > 0$ such that for all $\overline{u}(x) \epsilon \overline{F}$, $|u_i(x_1) - u_i(x_2)| \leq \epsilon$ for every $x_1, x_2 \epsilon I_{r'}$ satisfying $|x_1 - x_2| \leq \delta(\epsilon)$. Let C be the set of all functions $u(x) \epsilon X$ which satisfy (5-15) and are such that $|u_i(x_1) - u_i(x_2)| \leq \epsilon$ for every $x_1, x_2 \epsilon I_{r'}$ satisfying $|x_1 - x_2| \leq \delta(\epsilon)$. Then σ clearly transforms C into itself. Furthermore, C, being equicontinuous, is compact by Ascoli's Theorem [30, p.234] and is easily seen to be convex. Finally, since the $K_i[x, t, u_1, \ldots, u_n]$ are continuous on D and D is closed, they are uniformly continuous on D. Hence, σ is a continuous transformation of C into itself and is clearly single-valued. Thus, by applying Schauder's Theorem the following result is deduced.

Theorem 1: Let $f_i(x)$ be continuous on I_r and let $K_i[x, t, u_1, \ldots, u_n]$ on D. Then on $I_{r'}$ there exists at least one continuous solution of the integral equations given in (5-14) where $r' = \min(r, \rho/2M)$.

Of course, the solution of (5-14) can be extended to I_r or to the boundary of D by the standard argument.

By a straightforward application of the above theorem to the set of integral equations given in (5-7), the following important theorem is easily deduced.

Theorem 2: Suppose that $\Gamma(t_1, t_2)$ is of the form given by (3-4), that $q_i^{(1)}(t)$ and $p_i^{(1)}(t)$ exist and are continuous on $[0,T]$, that $J_{0,1}(t) > 0$ on $[0,T]$, and that there exists a nonnegative-definite matrix $[\Gamma_{ij}]$ such that $p_i(0) - \sum_{j=1}^{n} \Gamma_{ij} q_j(0)$ [illegible] 0 for all i. Then $\Gamma(t_1, t_2)$ is a covariance function on $[0,T]$ and there exists a *physically realizable* shaping filter on $[0,T]$ whose weighting function is of the form given in (2-19) where the $\rho_i(t)$ are continuous on $[0,T]$.

The only possible difficulty with the application of Theorem 1 to Theorem 2 is that it may be impossible to extend the solution to $[0,T]$ without leaving D. Since D itself [for (5-7)] can be increased in the u_i coordinate directions and still meet the requirements of Theorem 1 indefinitely, this implies that either the solution can be extended to $[0,T]$ or else the $\beta_i(t)$ become unbounded and, hence, discontinuous. *In the statement of Theorem 2 and in those to follow, it has been assumed that it is possible to extend the solution to* $[0,T]$. In any case, Theorem 2 holds on $[0,T']$ where $\rho/2M \leq T' \leq T$. Cases where the $\beta_i(t)$ are unbounded are not of great importance in engineering applications and, furthermore, computational problems arise in these cases anyway.

When $J_{0,1}(t) \equiv 0$ on $[0,T]$ but $J_{1,2}(t) > 0$ on $[0,T]$, then the following modified form of Theorem 2 holds.

Theorem 3: Suppose that $\Gamma(t_1,t_2)$ is of the form given by (3-4), that $q_i^{(2)}(t)$ and $p_i^{(2)}(t)$ exist and are continuous on $[0,T]$, that $J_{1,2}(t) > 0$ on $[0,T]$, and that there exists a nonnegative-definite matrix $[\Gamma_{ij}]$ such that $p_i(0) - \sum_{i=1}^{n} \Gamma_{ij} q_j(0) = 0$ and $p_i^{(1)}(0) - \sum_{i=1}^{n} \Gamma_{ij} q_j^{(1)}(0) = 0$ for all i. Then $\Gamma(t_1,t_2)$ is a covariance function on $[0,T]$ and there exists a *physically realizable* shaping filter on $[0,T]$ whose weighting function is of the form given in (2-19), the $\beta_i(t)$ are continuous on $[0,T]$, and $W(t,t) \equiv 0$ for $t \in [0,T]$.

The further modification of Theorem 2 when $J_{0,1}(t) = J_{1,2}(t) \equiv 0$ on $[0,T]$ but $J_{2,3}(t) > 0$ on $[0,T]$ is obvious. The case there $J_{0,1}(t) = 0$ for some $t \in [0,T]$ but not identically is not discussed, but satisfactory results could possibly be obtained by following up Lalesco's work [26].

Finally, suppose that in addition to satisfying the hypotheses of Theorem 2, the $q_i(t)$ and $p_i(t)$ have n-continuous derivatives on $[0,T]$ and the Wronskian of the $q_i(t)$ doesn't vanish on $[0,T]$. Then by successive differentiation of Equations (5-7) it follows that the $\beta_i(t)$ have $n - 1$ continuous derivatives on $[0,T]$. Hence, by the results in Section 2.4, it follows that the shaping filter can be characterized by a differential equation of the form given in (2-1) where the $a_i(t)$ and $b_j(t)$ are continuous on $[0,T]$.

5.5 THE QUESTION OF UNIQUENESS OF THE SHAPING FILTER

Having resolved the question of the existence of a *physically realizable* shaping filter, the question of the uniqueness of the shaping filter naturally arises. Examining (5-5), it is clear that there is no unique solution because if $W(t,\tau)$ is a solution, then $-W(t,\tau)$ is also a solution. Note that if $W(t,\tau)$ is the solution associated with the plus sign in (5-10), then $-W(t,\tau)$ is the solution associated with the minus sign. However, the question still remains as to whether the solution is unique, say, to within a multiplicative factor. The answer is again no because there may be more than one nonnegative-definite matrix $[\Gamma_{ij}]$ which meets the requirements of Theorem 2 and which lead to different solutions. As an example, consider the covariance function

$$\Gamma(t_1,t_2) = 4/3\, e^{-|t_1 - t_2|} - 5/12\, e^{-|t_1 - t_2|};\ t_1,\ t_2 \geq 0$$

The two matrices

$$[\Gamma'_{ij}] = \begin{bmatrix} 2 & -2/3 \\ -2/3 & 1/4 \end{bmatrix},\quad [\Gamma''_{ij}] = \begin{bmatrix} 8 & -20/3 \\ -20/3 & 25/4 \end{bmatrix} J_{0,1}(t) = 1$$

Both meet the requirements of Theorem 2. Furthermore, and as direct substitution into (5-7) shows, $\beta_1(t) = 2e^t$ and $\beta_2(t) = -e^{2t}$ is a solution on $[0,\infty]$ for $[\Gamma'_{ij}]$ while $\beta_1(t) = -4e^t$ and $\beta_2(t) = 5e^{2t}$ is a solution on $[0,\infty]$ for $[\Gamma''_{ij}]$. Hence, $W_1(t,\tau) = 2e^{-(t-\tau)} - e^{-2(t-\tau)}$ and $W_2(t,\tau) = -4e^{-(t-\tau)} + 5e^{-2(t-\tau)}$ are respectively the weighting functions of the physically realizable shaping filters for these matrices. Taking the Laplace transforms of $W_1(t-\tau)$ and $W_2(t-\tau)$ there results $G_1(S) = \dfrac{S+3}{(S+1)(S+2)}$ and $G_2(S) = \dfrac{S-3}{(S+1)(S+2)}$. It is interesting to note that the transfer function of the system associated with $[\Gamma'_{ij}]$ has its zero in the left-half plane while that associated with $[\Gamma''_{ij}]$ has its zero in the right-half plane. In light of this example, the question now arises as to whether the solution is unique if, say, the plus sign in (5-10) is used and a matrix $[\Gamma_{ij}]$ is specified which meets the requirements of Theorem 2. The answer this time is yes as the theorems proven below show.

Suppose that in addition to satisfying the hypotheses of Theorem 1, the $K_i[x_1,t,u_1,\ \ldots,\ u_n]/\partial u_j$ are continuous on D for all i,j. Then, since D is closed, they are bounded on D. By the mean value theorem, this implies that a Lipschitz condition is satisfied on D; i.e., that there exists an M_1 such $||K[x,t,u_1,\ \ldots,\ u_n] - K[x,t,v_1,\ \ldots,\ v_n]|| \leq M_1||u - v||$ where $||u - v|| = \sum_{i=1}^{n} |u_i - v_i|$, etc. Under this stronger assumption, the existence of a solution to (5-14) on I_r, can be established by successive approximations as follows. Let $u^0(x) = f(x)$ and $u^{j+1}(x)$

$$u_i^{j+1}(x) = f_i(x) + \int_0^x K_i[x,t,u_1^j(t), \ldots, u_n^j(t)]\, dt \tag{5-16}$$

Then $u^{j+1}(x)\epsilon D$ and is continuous for ϵI_r, as is easily established by induction. Furthermore

$$||u^{j+1}(x) - u^j(x)|| \leq M_1 \int_0^x ||u^j(t) - u^{j-1}(t)||dt \tag{5-17}$$

and

$$||u^1(x) - u^0(x)|| \leq \int_0^x ||K[x,t,f_1(t), \ldots, f_n(t)]||\, dt \tag{5-18}$$
$$\leq nM\, x$$

By iteration, using (5-17) there results

$$||u^{j+1}(x) - u^j(x)|| \leq \frac{nM\, M_1^j\, x^{j+1}}{(j+1)!} \tag{5-19}$$

Hence, the series $\sum_{j=0}^{\infty} ||u^{j+1}(x) - u^j(x)||$ converges uniformly on I_r, which in turn implies that $u^j(x)$ converges uniformly on I_r, to a continuous vector function $u(x)$ which satisfies (5-14). The uniqueness $u(x)$ can be established as follows. Suppose $v(x)$ is another solution of (5-14) in D on I_r,. Then, it follows that (making use of the Lipschitz condition)

$$||u^{j+1}(x) - v(x)|| \leq M_1 \int_0^x ||u^j(t) - v(t)||\, dt \tag{5-20}$$

Again, using the fact that $||u^0(x) - v(x)|| \leq nm\, x$, by iteration

$$||u^{j+1}(x) - v(x)|| \leq \frac{nM\, M_1^j\, x^{j+1}}{(j+1)!} \tag{5-21}$$

which on letting $j \to \infty$ implies that $||u(x) - v(x)|| \leq 0$. Hence, $u(x) = v(x)$ on I_r, and the solution is unique. This proves the following theorem.

Theorem 4: If the hypotheses of Theorem 1 are satisfied and in addition $K_i[x,t,u_1, \ldots, u_n]/\partial u_j$ are continuous on D, then there exists a continuous solution of (5-14) on I_r, which can be found by successive approximations and the solution is unique.

As before, the solution can be extended to I_r or to the boundary of D by the standard argument. Application of Theorem 4 to (5-7) yields immediately the following important theorem.

Theorem 5: If the hypotheses of Theorem 2 or Theorem 3 are satisfied, then a *physically realizable* shaping filter exists on $[0,T]$ and, if the sign in (5-10) is chosen and the matrix $[\Gamma_{ij}]$ specified, the shaping filter is unique.

This concludes the discussion of existence and uniqueness of *physically realizable* shaping filter.

CHAPTER 6

COMPUTATIONAL ASPECTS AND APPLICATIONS

6.1 COMPUTATIONAL ASPECTS

Since the engineer is usually interested in synthesizing shaping filters for use in applications, he is faced with the problem of actually finding the weighting functions and/or differential equations characterizing them. Except for certain special cases such as stationary processes over the interval $(-\infty,\infty)$ and those discussed in Chapter 4, the determination of the $\beta_i(t)$ (or the coefficients of the corresponding differential equation) analytically appears to be a very difficult, if not impossible, task. Thus, one is naturally led to the consideration of computational methods. A few remarks on this aspect of the problem are given below. One of the practical justifications of the work in Chapter 5 which is apparent here is that it establishes the existence of a solution and its uniqueness properties at the outset of the problem, thereby guaranteeing that one is not trying to compute something that does not exist.*

In view of Theorems 4 and 5, one of the immediate methods which come to mind for computing the $\beta_i(t)$ from (5-7) is that of successive approximations. This, of course, can be done on either a digital or an analog computer. Standard references on numerical methods such as Hildebrand [32] discuss the problem from the standpoint of digital computation and it will not be discussed further here. In a recent article, Tomovic and Parezanovic [33] have investigated the use of repetitive analog computers for solving integral equations by successive approximations. The interested reader is referred to this article and those referenced therein.

Since the integral Equations (5-7) are of the Volterra type, they can also be integrated directly either digitally or on an analog computer in much the same manner as differential equations for one point boundary value problems are integrated. This is perhaps a better over-all computational procedure than successive approximations. Finally, Equations (5-5) can also be solved on an analog computer by implicit methods.

While a great deal of effort could be spent on developing optimal computational algorithms for integral equations of the type given in (5-5) and (5-7), it appears that the standard techniques mentioned above are adequate considering the need.

---

*Of course, the work in Chapter 5 is justifiable in its own right because of, among other things, the insight it gives into the structure of certain classes of stochastic processes.

6.2 APPLICATIONS

The two main areas of application of shaping filters are to the analysis (usually on analog computers) of the effects of noise on linear systems and to the design of linear least-squares, smoothing and predicting filters. While these applications are well known, a short presentation of them will be given in the interest of completeness.

The first application is to the problem of finding the variance of the output of a linear system when the input is a stochastic process whose covariance function is known. This problem reduces to the computation of an integral of the form

$$\sigma^2(t) = \int_0^t d\tau_1 \int_0^t d\tau_2 \; W(t,\tau_1)W(t,\tau_2)\Gamma(\tau_1, \tau_2) \tag{6-1}$$

where $W(t,\tau)$ is the weighting function of the system, $\Gamma(\tau_1,\tau_2)$ is the covariance function of the input process, and $\sigma^2(t)$ is the variance of the output process as a function of time. When a shaping filter exists for the input process, then

$$\Gamma(\tau_1,\tau_2) = \int_0^{\tau_1} d\theta_1 \int_0^{\tau_2} d\theta_2 \; W_s(\tau_1,\theta_1)W_s(\tau_2,\theta_2)\delta(\theta_1 - \theta_2) + \\ + \sum_{i=1}^{n} \sum_{j=1}^{n} \Gamma_{ij}\, q_i(\tau_1)\, q_j(\tau_2) \tag{6-2}$$

Substitution of (6-2) into (6-1), interchanging the order of the integrations, and integrating out the δ function yields

$$\sigma^2(t) = \int_0^t d\theta \left[\int_\theta^t d\tau_1 \; W(t,\tau_1)\, W_s(\tau_1,\theta)\right]^2 \\ + \sum_{i=1}^{n} \sum_{j=1}^{n} \Gamma_{ij} \int_0^t d\tau_1 \; W(t,\tau_1)\, q_i(\tau_1) \; \times \\ \times \int_0^t d\tau_2 \; W(t,\tau_2)\, q_j(\tau_2) \tag{6-3}$$

The term in square brackets in (6-3) represents the weighting function, $\overline{W}(t,\theta)$, of the cascade of the system and the shaping filter. The computation of the first integral in (6-3) is easily carried out on an analog computer by the method of adjoint systems described in Laning and Battin [20] when the system is characterized by a finite-order linear differential equation, the covariance function is separable, and the Wronskian of the $q_i(t)$ exists and doesn't vanish on the interval of interest. The other terms in (6-3) can obviously be computed separately with part of the same analog setup. The restriction to separable covariance functions and the nonvanishing of the Wronskian of the $q_i(t)$ allows the shaping filter to be simulated as shown in Figure 3, where the $a_i(t)$ are given by (2-31) and the $F_i(t)$ are given by [making use of (2-26)]

$$F_i(t) = \sum_{j=1}^{n} q_j^{(i-1)}(t)\, \beta_1(t) \qquad (6\text{-}4)$$

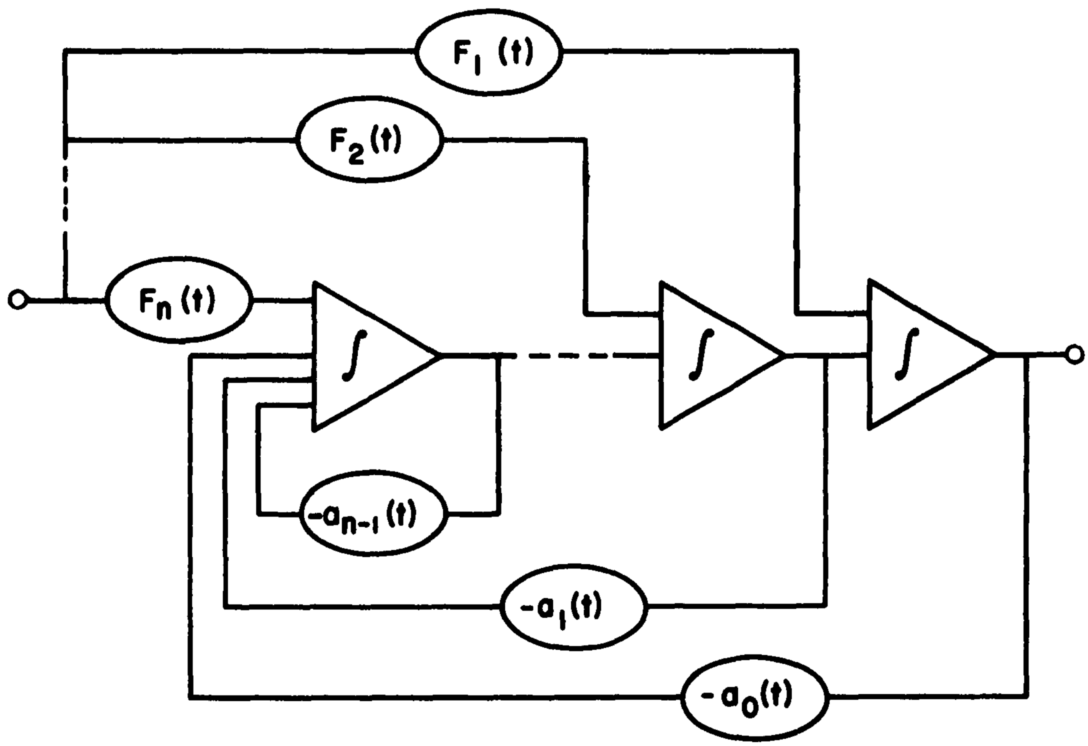

FIGURE 3

Note that no differentiability is required of the $\beta_i(t)$ for this form of simulation which in turn requires only continuity of first derivatives of $p_i(t)$ to guarantee continuity of the $\beta_i(t)$ and, hence, continuity of the $F_i(t)$.

The other area of application of shaping filters is to the design of linear, least-squares, smoothing and predicting filters. The usual formulation of this problem leads to the Wiener-Hopf integral equation

$$\Gamma_{DI}(t_1, t_2) = \int_0^{t_1} d\theta\, W(t_1, \theta)\, \Gamma_{II}(\theta, t_2) \quad ; \quad t_1 \geq t_2 \geq 0 \qquad (6\text{-}5)$$

where $\Gamma_{DI}(t_1, t_2)$ is the covariance function of the desired *signal* at the present time t_1 and the input at time t_2, $\Gamma_{II}(\theta, t_2)$ is the covariance function of the input at time θ and time t_2, and $W(t_1, \theta)$ is the weighting function of the desired least-squares filter. When the input is a *white noise* process, then $\Gamma_{II}(\theta, t_2) = \delta(\theta - t_2)$ and (6-5) can be solved immediately, yielding

$$W(t_1, t_2) = \begin{cases} \Gamma_{DI}(t_1, t_2) & ; \ t_1 \geq t_2 \geq 0 \\ 0 & ; \ t_1 < t_2 \end{cases} \tag{6-6}$$

The basic approach used above can still be used even when the input is not a *white noise* process providing there exists a physically realizable linear system (called an inverse shaping filter) whose response to the input will be a *white noise* process. In this case, following the Bode-Shannon idea, the input is first operated on by the inverse shaping filter yielding a *white noise* process. Treating the output of the inverse shaping filter as a new input process and applying the result of the previous paragraph one obtains

$$W_1(t_1, t_2) = \begin{cases} \Gamma_{DI'}(t_1, t_2) & ; \ t_1 \geq t_2 \geq 0 \\ 0 & ; \ t_1 < t_2 \end{cases} \tag{6-7}$$

where

$$\Gamma_{DI'}(t_1, t_2) = \int_0^{t_2} d\tau \ W_s^{-1}(t_2, \tau) \ \Gamma_{DI}(t_1, \tau) \tag{6-8}$$

Here $W_s^{-1}(t_2, \tau)$ denotes the weighting function of the inverse shaping filter.* Hence, the weighting function of the least-squares filter is given by

$$W(t_1, t_2) = \int_{t_2}^{t_1} d\tau \ W_1(t_1, \tau) [W_S^{-1}(\tau, t_2)]^a \tag{6-9}$$

For cases where there exists a shaping filter which is characterized by a finite-order linear differential equation, the differential equation of the inverse shaping filter is immediately found by interchanging the role of input and output. Also, in this case, the integral in (6-8) is easily evaluated on an analog computer. When $\Gamma_{DI}(t_1, t_2)$ is separable, then $W(t_1, t_2)$ is separable and, assuming the requisite differentiability, the least-squares

*If one is operating over an infinite interval, then the inverse shaping filter must be stable in the usual sense. For a finite interval, stability in the usual sense loses its meaning and importance.

filter is characterized by a finite-order linear differential equation which is easily simulated (or built) from analog components.

From the above it is clear that if the differential equation of the shaping filter is known, then the solution of the least-squares filtering problem is greatly simplified. For further discussion see Darlington [9].

Kalman and Bucy [34] have given an alternate solution to the least-squares filtering problem assuming that the shaping filter for the *signal* is known and the noise is *white noise*.* As before, the solution makes use of an explicit knowledge of the differential equation of a shaping filter.

It should be noted that, if random initial conditions are required on the shaping filter, then, the approach discussed above must be modified. Kalman and Bucy claim that their results hold for this case without modification.

---

*There are processes for which a shaping filter does not exist as, for example, stationary processes with nonabsolutely continuous spectral distribution functions. Hence, Kalman has not solved all the problems as he sometimes claims.

REFERENCES

1. Wold, H., *A Study in the Analysis of Stationary Time Series,* Uppsala, 1938.
2. Kolmogorov, A., "Sur l' Interpolation et Extrapolation des Suites Stationnaires", *C. R. Acad. Sci.*, Paris 208, pp. 2043-45, 1939.
3. Kolmogorov, A., "Stationary Sequences in Hilbert Space", (Russian), *Bull. Math. Univ.*, Moscow 2, No. 6, 40 pp., 1941.
4. Kolmogorov, A., "Interpolation und Extrapolation von Stationaren Zufalligen Folgen", *Bull. Acad. Sci.*, U.S.S.R. Ser. Math. 5, 3-14, 1941.
5. Wiener, N., *Extrapolation, Interpolation, and Smoothing of Stationary Time Series,* John Wiley and Sons, Inc., New York, 1949.
6. Hanner, O., "Deterministic and Nondeterministic Stationary Random Processes", *Ark. Mat. 1,* pp. 161-77, 1949.
7. Karhunen, K., "Uber die Struktur Stationarer Zufalliger Funktionen", *Ark Mat. 1,* pp. 141-60, 1960.
8. Bode, H. and C. Shannon, "A Simplified Derivation of Linear Least-Square Smoothing and Prediction Theory", *Proc. IRE* 38, pp. 417-26, April 1950.
9. Darlington, S., "Nonstationary Smoothing and Prediction Using Network Theory Concepts", *Transactions IRE,* I T-5, pp. 1-11, May 1959.
10. Batkov, A., "Generalization of the Shaping Filter Method to Include Nonstationary Random Processes", *Automation and Remote Control,* Vol. 20, No. 8, August 1959, pp. 1049-62.
11. Sondhi, M.M., and T. J. Higgins, *A Procedure for Synthesizing Linear Time-Varying Shaping Filters for Generating Nonstationary Random Outputs,* presented at the NACC, 1959, Dallas.
12. Leonov, Y., "The Problem of Shaping Filters and Optimal Linear Systems", *Automation and Remote Control,* Vol. 21, No. 6, June 1960, pp. 467-71.
13. Doob, J., *Stochastic Processes,* John Wiley and Sons, Inc., New York, 1953.
14. Grenander, U., and M. Rosenblatt, *Statistical Analysis of Stationary Time Series,* John Wiley and Sons, Inc., New York, 1957.
15. Levy, P., "The Integration of a Nonlinear Integral Equation", *C. R. Acad. Sci.,* 242, 10, 1956.
16. Pugachev, V., *Theory of Random Functions and Their Application to Automatic Control Problems,* (Russian), Gostekhizdat, 1957.
17. Pugachev, V., "Integral Canonical Representation of Random Functions and Their Application in Deriving Optimal Linear Systems", *Automation and Remote Control,* Vol. 18, No. 11, November 1957, pp. 1017-31.
18. Natanson, I., *Theory of Functions of a Real Variable,* Frederick Ungar Publishing Co., New York, 1955.
19. Ince, E., *Ordinary Differential Equations,* Dover Publications, Inc., New York, 1956.

20. Laning, J., and R. Battin, *Random Processes in Automatic Control*, McGraw-Hill Book Company, Inc., New York, 1956.

21. Coddington, E., and N. Levinson, *Theory of Ordinary Differential Equations*, McGraw-Hill Book Company, Inc., New York, 1955.

22. Borskii, V., "On the Properties of the Impulsive Response Function of Systems with Variable Parameters", *Automation and Remote Control*, Vol. 20, No. 12, December 1959, pp. 1544-49.

23. Loeve, M., *Probability Theory*, D. Van Nostrand Company, Inc., New York, 1960.

24. Halmos, P., *Introduction to Hilbert Space*, Chelsea Publishing Company, New York, 1957.

25. Tricomi, F., *Integral Equations*, Interscience Publishers, Inc., New York, 1957.

26. Lalesco, T., *Introduction a la Theorie des Equations Integrales*, Gauthier-Villars, Paris, 1922.

27. Sato, T., "Sur l'Equation Integrale Non Linear de Volterra", *Compositio Mathematica*, 11, 1953, pp. 271-90.

28. Lefschetz, S., *Introduction to Topology*, Princeton University Press, 1949.

29. Berge, C., *Espaces Topologiques*, Dunod, Paris, 1959.

30. Kelley, J., *General Topology*, D. Van Nostrand Company, Inc., 1955.

31. Bellman, R., *Stability Theory of Differential Equations*, McGraw-Hill Book Company, Inc., 1953.

32. Hildebrand, F., *Introduction to Numerical Analysis*, McGraw-Hill Book Company, Inc., 1956.

33. Tomovic, R., and N. Parezanovic, "Solving Integral Equations on a Repetitive Differential Analyzer", *IRE Trans. on Electronic Computers*, Vol. EC-9, No. 4, December 1960, pp. 503-507.

34. Kalman, R. and R. Bucy, *New Results in Filtering and Prediction Theory*, presented at Joint Automatic Control Conference, M.I.T., 1960.

APPENDIX

Suppose the $q_i(t)$ are linearly independent on I in the sense that for all finite n there exists no set of constants a_i not all zero such that $\sum_{i=1}^{n} a_i\, q_i(t) = 0$ almost everywhere on I and suppose also that $\int_I q_i^2(t)\, dt < \infty$ for all i. Let $\mathfrak{N}$ be the closed linear manifold generated by all of the $q_i(t)$; i.e., the set of all functions of the form $\sum_{i=1}^{m} a_i\, q_i(t)$ which are square integrable over I or limits in the mean of such sums where the a_i are arbitrary constants and m is an arbitrary integer, $1 \leq m < \infty$. Let $\mathfrak{M}_j$ be the closed linear manifold generated by all of the $q_i(t)$ except $q_j(t)$. Then $\mathfrak{M}_j \subset \mathfrak{N}$ and by a well known result [24,p23] there exists a function $f_j(t) \epsilon \mathfrak{N}$, $\int_I f_j^2(t)\, dt > 0$, such that $f_j(t) \perp \mathfrak{M}_j$; i.e., such that $\int_I f_j(t) \times q(t)\, dt = 0$ for all $q(t) \epsilon \mathfrak{M}_j$. Since $q_i(t) \epsilon \mathfrak{M}_j$ for all $i \neq j$, this shows that there exists a function $f_j(t)$ such that $\int_I f_j(t)\, q_i(t)\, dt = 0$ for $i \neq j$. Clearly $\int_I f_j^2(t)\, dt > 0$ and $f_j(t) \perp \mathfrak{M}_j$ implies $|\int_I f_j(t)\, q_i(t)\, dt| > 0$, and thus, by suitably norming, this establishes the existence of a function $f_j(t)$ such that $\int_I q_i(t)\, f_j(t)\, dt = \delta_{ij}$. Since this can be done for every $q_j(t)$, this establishes the existence of a set of functions $f_i(t)$ which together with the $q_i(t)$ form a set of biorthonormal functions on I. Note that since $f_i(t) \epsilon \mathfrak{N}$ for all i, $\int_I f_i^2(t)\, dt < \infty$ for all i. Finally, the $f_i(t)$ are obviously linearly independent on I in the same sense as the $q_i(t)$.